MANITOBA WRITING ASSESSMENT PROGRAM 1982

PRELIMINARY REPORT: TEST DATA

A REPORT OF THE
CURRICULUM DEVELOPMENT AND IMPLEMENTATION BRANCH
DEPARTMENT OF EDUCATION
PROVINCE OF MANITOBA

371.26097127
Manitoba Writing Assessment Program, 1982
Preliminary Report.

ISBN 0-86497-093-5
1. Writing education--Manitoba--Ability testing. 2. Educational tests and measurements--Manitoba. I. Manitoba. Dept. of Education. Curriculum Development and Implementation Branch

Winnipeg
October, 1982

THE 1982 MANITOBA WRITING ASSESSMENT PROGRAM

The Assessment required the development of writing tests at grade levels 3, 6, 9, and 12. These tests differed from the ones developed for the 1979 Writing Assessment in several significant ways, including the setting of two writing tasks at the grade 3, 6, and 9 levels and a planning and a writing task at the grade 12 level. However, certain aspects of the testing and scoring procedures were designed to enable comparisons of student performance from 1979 to 1982.

Two reports are being produced on the results of the Assessment: the Preliminary Report which contains detailed results of the performance of students in 1982, and the Final Report which will contain the conclusions and recommendations based on these results and, in part, on comparisons with the 1979 results.

A major part of the design and development of the testing materials and scoring schemes, the actual training of teachers for scoring of the provincial sample, and the design of the preliminary and final reports was due to the contractor: Dr. Stanley Straw of the University of Manitoba.

PREFACE

The Preliminary Report is the first of two reports on the findings of the Writing Assessment Program 1982. It contains a description of the assessment program and results of student performance on the grade 3, 6, 9, and 12 writing tests. Conclusions and recommendations will be presented in the Final Report which will be distributed to schools later in the school year.

The purpose of the Preliminary Report is to provide detailed information on student performance for use by classroom teachers as part of their instructional program. All data is based on the provincial sample as a whole and no individuals, schools, or divisions are identified. Teachers are encouraged to review the results of their own students in the light of those of the province. The test questions and booklets may be used, in whole or in part, as descriptive tools for instructional purposes.

In this report no attempt has been made to interpret the results. Comparisons with previous performance and conclusions and recommendations will be made after further analyses and review of the results.

For additional information contact:

Henry Schulz
Curriculum Development and Implementation Branch
Department of Education
223-1200 Portage Avenue
Winnipeg, Manitoba
R3G OT5

Telephone: 786-0144 (Long distance toll-free WATS LINE 1-800-262-8848, Extension 130).

COMMITTEES AND GROUPS ASSOCIATED WITH THE ASSESSMENT

Joint Committee on Evaluation

Manitoba Teachers' Society: John Gisiger
Tina Kehler (to September, 1982)
Judy Kruger (from September, 1982)
Manitoba Association of School Trustees: Michael Buyachok
Roy White
Manitoba Association of School Superintendents:
Robert Cross, Chairperson
University of Manitoba, Faculty of Education:
Winston Rampaul (to September, 1982)
Barbara Nelson (from September, 1982)
Brandon University, Faculty of Education: Pandy Halamandaris (to June, 1982)
James Hardy

Manitoba Department of Education,
Curriculum Development and Implementation Branch:
Stan Bullock (to September, 1982)
Ken Charlesworth (from September, 1982)
Henry Schulz
Bureau de l'Education Francaise:
Gilbert Rosset (to June, 1982)
Raymond Theberge (from June, 1982)

Consultants

Manitoba Department of Education,
Research Branch: Mike Yakimishyn
Curriculum Development and Implementation Branch:
Norman Mayer
Conrad Sigurdson

Recorder

Manitoba Department of Education,
Curriculum Development and Implementation Branch:
Barbara Brooks

Writing Technical Advisory Committee

Henry Schulz (Chairperson), Curriculum Development and Implementation Branch, Manitoba Department of Education
Cheryl Billinkoff, Arthur E. Wright Community School, Seven Oaks School Division #10
Steve Britton, Crocus Plains Regional Secondary School, Brandon School Division #40
Lorie Camden, Elmwood School, Rhineland School Division #18
Ron Cancilla, Chief Peguis Junior High, River East School Division #9
Eileen Fast, Parc La Salle School, Seine River School Division #14
Shirley Hicks, Leaf Rapids Ed Center, Leaf Rapids School District #2460
Margaret Johnston, McAuley School, Fort La Bosse School Division #41
Ray Lavery, Garden City Collegiate, Seven Oaks School Division #10
Heather Lytwyn, Hapnot Collegiate, Flin Flon School Division #46
David Oster, Arthur E. Wright Community School, Seven Oaks School Division #10
Kristine Perlmutter, Graduate Student, University of Manitoba
Michael Zawaly, Happy Thought School, Lord Selkirk School Division #11

Consultants

Manitoba Department of Education, Curriculum Development and Implementation Branch:
JoAnne Beavis
Don Bewell
Dorothy Brummit
Winston Layne

Contractor

Dr. Stanley Straw, University of Manitoba

COMMITTEES AND GROUPS ASSOCIATED WITH THE ASSESSMENT

Scorers for the Provincial Sample

Grade 3

Barbara Brooks, Curriculum Development and Implementation Branch, Manitoba Department of Education
Claudia Chomichuk, Marion School, St. Boniface School Division #4
Mary Chubey, Shevchenko Primary School, Boundary School Division #16
Rosaline Fedorkie, Fort Rouge School, Winnipeg School Division #1
Christine Frank, Oakbank Elementary School, Transcona-Springfield School Division #12
Bruce Holden, Hamiota Elementary School, Birdtail River School Division #38
Laraine Hunt, Maple Leaf Elementary, Western School Division #47
Gail Johnson, Binscarth Elementary School, Pelly Trail School Division #37
Thelma Jamieson, Forest Park School, Seven Oaks School Division #10
Doris Kristjanson, Ralph Maybank School, Fort Garry School Division #5
Leona Mackedenski, Onanole School, Rolling River School Division #39
Freda Mostoway, Major Pratt School, Pelly Trail School Division #37
Wendy Ohryn, Kelsey Elementary, Kelsey School Division #45
Anne Plohman, Ashern Central School, Lakeshore School Division #23
Pat Robson, Lincoln School, St. James-Assiniboia School Division #2
Earl Stopera, Waldheim Elementary, White Horse Plain School Division #20
Patricia Williams, O'Kelly School, Camp Shilo School District #2316
Charlene Wilson, Laura Secord School, Winnipeg School Division #1
Merle Zbrog, Dieppe School, Assiniboine South School Division #3

Grade 6

Judith Davies, Curriculum Development and Implementation, Manitoba Department of Education

Grade 6 (Cont'd.)

Doug Barron, Manitou Elementary, Pembina Valley School Division #27
Bob Bastable, Mapleton School, Lord Selkirk School Division #11
Debra Black, Cecil Rhodes School, Winnipeg School Division #1
Anna Bowman, Substitute Teacher, Winnipeg School Division #1
Greg Bryan, Onanole Elementary School, Rolling River School Division #39
Pamela Davies, Elwick Community School, Seven Oaks School Division #10
Dan De Baets, General Byng School, Fort Garry School Division #5
Marnie Erb, J. A. Cuddy Elementary, Morris-MacDonald School Division #19
Al Friesen, Windsor School, St. Vital School Division #6
Debbie Goytan, Birtle Elementary, Birdtail River School Division #38
Wayne Hominick, Ecole St. Adolphe, Seine River School Division #14
Phyllis Hunter, Maple Leaf Elementary, Western School Division #47
Carole Locht, Woodhaven School, St. James-Assiniboia School Division #2
Leslie Maxted, Hazel M. Kellington School, Beautiful Plains School Division #31
David B. Oster, Arthur E. Wright Community School, Seven Oaks School Division #10
Iris Reynolds, Cecil Rhodes School #2, Winnipeg School Division #1
Robin Robinson, Binscarth School, Pelly Trail School Division #37
Menno Schulz, Ralph Maybank School, Fort Garry School Division #5
Andy Tardiff, Marion School, St. Boniface School Division #4

Grade 9

Henry Schulz, Curriculum Development and Implementation, Manitoba Department of Education,
Loretta Basiuk, Edward Schreyer School, Agassiz School Division #13

Grade 9 (Cont'd.)

John D. Bender, MacGregor Collegiate, Pine Creek School Division #30
Cindy Bigford, Acadia Junior High, Fort Garry School Division #5
F. Donald Clark, J. M. Young School, Beautiful Plains School Division #31
Arlene J. Day, Rossendale School, Pine Creek School Division #30
Morris P. Hykaway, St. Pierre Collegiate, Red River School Division #17
Klaus H. Kelm, Alf Cuthbert School, Lakeshore School Division #23
Don Kyliuk, Beliveau Junior High, St. Boniface School Division #4
Olga K. Kowalchuk, Morse Place Junior High, River East School Division #9
Wesley Lasuik, Swan River Junior High, Swan Valley School Division #35
Paul Leger, Varennes School, St. Vital School Division #6
Helen Leggitt, Acadia Junior High, Fort Garry School Division #5
Edwin Lenzmann, Grant Park High School, Winnipeg School Division #1
William A. McCallum, MacKenzie Junior High, Dauphin-Ochre School Division #33
Charlotte E. L. Milan, Margaret Barbour Collegiate Institute, Kelsey School Division #45
Sandy Millen, River Heights Junior High, Winnipeg School Division #1
Kristine Perlmutter, Graduate Student, Univ. of Manitoba
John G. Reimer, Landmark Collegiate, Hanover School Division #15
Norman Sanders, St. John's High School, Winnipeg School Division #1
Wayne Selby, Ecole Precieux Sang, Norwood School Division #8
Ken Sheppard, Charleswood Collegiate, Assiniboine South School Division #3
Gerald Shewchuk, MacKenzie Junior High, Dauphin-Ochre School Division #33
Cathy Andrayko Smith, Hedges Junior High, St. James-Assiniboia School Division #2
Jan Starling, Substitute Teacher, Brandon School Division #40
Donald W. J. Whyte, Prince Charles School, Portage la Prairie School Division #24

Grade 12

Ralph Wagner, Curriculum Development and Implementation, Manitoba Department of Education
John Blewett, Argyle School, Winnipeg School Division #1
Roy Bowkett, Tuxedo-Shaftsbury High School, Assiniboine South School Division #3
Beverley Brown, Rossburn Collegiate, Pelly Trail School Division #37
Joanne Carels, Treherne Collegiate, Tiger Hills School Division #29
Donald E. Cronk, Dauphin Regional Comprehensive Secondary School, Dauphin-Ochre School Division #33
Patricia Feir, Shevchenko School, Boundary School Division #16
Norman W. Fullerton, Fort Richmond Collegiate, Fort Garry School Division #5
Helen Levi, Glenboro School, Tiger Hills School Division #29
J. Lovelace, Ste. Anne School, Seine River School Division #14
Donald W. Luy, Edward Schreyer School, Agassiz School Division #13
Edward Mann, Vincent Massey Collegiate, Fort Garry School Division #5
Helen Mazur, Springfield Collegiate, Transcona-Springfield School Division #12
Phil McBurney, Kildonan East Reg. Sec. School, River East School Division #9
Jeff Milan, Margaret Barbour Collegiate Institute, Kelsey School Division #45
Jim Pehura, Portage Collegiate Institute, Portage la Prairie School Division #24
Tammy Trowsky, Elkhorn School, Fort la Bosse School Division #41
Marion Walkey, Hapnot Collegiate, Flin Flon School Division #46
Alfred Wiebe, West Kildonan Collegiate, Seven Oaks School Division #10

ACKNOWLEDGEMENTS

This assessment would not have been possible without assistance from the many people involved in development of the tests and production of reports, and co-operation of the students who wrote the tests and the teachers of these students. Special mention must be made of a number of groups:

- The contractor who gave professional assistance to all aspects of the testing and reporting.
- The Joint Committee on Evaluation which provided guidance throughout the program.
- The Writing Technical Advisory Committee which advised on test production and analysis of results.
- The teachers and others who participated in the review of objectives.
- The school administrators, teachers, and students that were involved in the pilot testing.
- The divisions and schools that released their teachers to assist in the program.
- The staff of the Curriculum Development and Implementation Branch and other Branches of the Department of Education who provided necessary support throughout the program.

A WORD OF CAUTION ABOUT THE ASSESSMENT PROGRAM

All measurements of student performance are subject to errors of measurement. These errors can occur in a number of ways: the specification of skills and objectives, the selection of materials and of test questions, the interpretation of questions by students, the scoring of student responses, the presentation of results, and the interpretation of these results. In the assessment program, effort has been made to minimize the error; nevertheless, the results should be regarded as estimates of the performance of students. Individual student diagnosis or grading on the basis of the tests should be attempted only in conjunction with other tests and assessments. Furthermore, the intent of the assessments is to obtain estimates of the performance of Manitoba students in general, and not of particular groups.

In the interpretation of results it is not possible to attribute performance strengths and weaknesses to specific aspects of curriculum and instruction. In addition, a variety of societal factors must be considered. The data presented in this report provide initial estimates of the performance of Manitoba students. One part of the interpretation process will be a comparison of selected aspects of the 1982 writing results to those obtained in the 1979 provincial Writing Assessment Program.

TABLE OF CONTENTS

Page

PREFACE. iii

COMMITTEES AND GROUPS ASSOCIATED WITH THE ASSESSMENT . iv

SCORERS FOR THE PROVINCIAL SAMPLE. v

ACKNOWLEDGEMENTS vii

A WORD OF CAUTION ABOUT THE ASSESSMENT PROGRAM vii

CHAPTER

1. BACKGROUND TO THE ASSESSMENT 1

- STEPS IN THE ASSESSMENT PROGRAM 1
- ASSESSMENT IN WRITING 3
- PROCEDURES FOR COMPARISON WITH THE 1979 ASSESSMENT 6
- REFERENCES. 6

2. GRADE 3 TEST RESULTS 7

- WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES. 8
- WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 22
- SPELLING. 32
- SENTENCE WRITING. 34

3. GRADE 6 TEST RESULTS 41

- WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES. 42
- WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS. 56
- SPELLING. 66
- SENTENCE WRITING. 68

Page

4. GRADE 9 TEST RESULTS. 75

- WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 76
- WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 90
- SPELLING. 100
- SENTENCE WRITING. 102

5. GRADE 12 TEST: RESULTS FOR ENGLISH 300 STUDENTS 107

- WRITTEN PRODUCT: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 108
- WRITTEN PRODUCT: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 116
- ESSAY PLANNING. 122
- SPELLING. 126
- SENTENCE WRITING. 128

6. GRADE 12 TEST: RESULTS FOR ENGLISH 301 STUDENTS 133

- WRITTEN PRODUCT: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 134
- WRITTEN PRODUCT: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 142
- ESSAY PLANNING. 148
- SPELLING. 152
- SENTENCE WRITING. 154

CHAPTER 1

BACKGROUND TO THE ASSESSMENT

The 1982 Writing Assessment Program is part of the overall Manitoba student assessment program as recommended by the Joint Committee on Evaluation and approved by the Minister of Education. The purposes for conducting assessment programs are as follows:

To provide benchmark indicators about the level of student achievement in the Province of Manitoba.
To obtain data on student achievement that will assist in curriculum and program improvement at the provincial and local level.
To assist school divisions in student and system evaluation.
To help teachers improve their student evaluation skills.

As part of the overall process, comparisons of student performance with previous performance are conducted wherever this is possible.

The writing tests were designed to describe student writing performance at grade levels 3, 6, 9, and 12. For the provincial survey, tests were administered to a sample of students selected from public schools throughout Manitoba.

This, the Preliminary Report contains an overview of the procedures which were followed and a detailed presentation of the student data collected for the Writing Assessment Program 1982.

The data are presented in tabular form and are accompanied by a minimum of discussion, thereby providing student information for use by classroom teachers. All data are reported on a provincial basis only with no individuals, schools, or school divisions being identified.

In the Preliminary Report no attempt has been made to interpret any of the test results. Following distribution of this report to schools, meetings of teachers and others will be held. Interpretive information, comparisons with the 1979 Writing Assessment, and conclusions will be contained in a Final Report together with more comprehensive findings and recommendations.

A teacher survey is being conducted in conjunction with the assessment in order to provide teacher information that would complement the student achievement data. Results of the teacher survey will be included in the Final Report.

STEPS IN THE ASSESSMENT PROGRAM

The procedures of the Writing Assessment are similar to those followed for previous Provincial assessments. Following the selection of Writing as the area for assessment, a Writing Technical Advisory Committee (Writing TAC) and a contractor were established to assist in the actual production of the tests and reporting of the results.

The Writing TAC consisted of twelve teachers with experience in language arts, one member of the Curriculum Development and Implementation Branch, four members of the Curriculum Services Branch of the Department of Education, and the contractor. The member from the Curriculum Development and Implementation Branch served as chairperson of the TAC.

The contractor is an education professor, with the University of Manitoba, experienced in the teaching and assessment of language arts.

The major steps involved in the Assessment Program are outlined below.

Step One: Identification of Objectives

An initial list of objectives for writing and a model by which they could be assessed was produced by the contractor. These were reviewed by the TAC and prepared for discussion with practicing language arts teachers.

Step Two: Selection of Objectives

Teachers of language arts from districts and divisions throughout Manitoba supplemented and rated the importance of each objective for instruction and assessment at the particular grade level.

Regional meetings were held in various locations in Manitoba for this purpose. The data collected at the meetings were compiled and used in the selection of objectives for the tests.

Step Three: Production of Test Items

The TAC recommended the format and the specifications for the tests based on the input from teachers at the regional meetings. The contractor, with the assistance of the TAC, produced test items according to the specifications. All test items were reviewed and modified by the TAC before inclusion in the pilot tests. This was done to ensure accuracy of the wording, appropriate readability, and technical quality of the items.

Step Four: Pilot Testing

The pilot tests were administered to several classes of students at each of the four grade levels. The test results were analyzed to determine the adequacy of test items. Items were modified or deleted as necessary by the TAC and rewritten by the contractor.

The final versions of the tests were reviewed by the TAC.

Step Five: Provincial Administration of Tests

The tests were administered on the morning of Wednesday, May 19, 1982, to students in grades 3, 6, 9, and 12 (those in English 300 and 301 programs).

Approximately 10 percent of the students in these grades from Manitoba public schools were randomly designated for the provincial sample. Table 1 indicates the provincial sample of tests at each grade level. The actual sample of student tests analyzed was approximately 88 percent of the designated sample. The sample loss was due primarily to student absence and transfer and varied somewhat according to grade level.

TABLE 1

PROVINCIAL SAMPLE OF TESTS

Grade	Number of Students Eligible	Students in the Sample	
		Number	Percentage
3	14,649	1,333	9.10
6	15,419	1,430	9.27
9	13,897	1,197	8.61
12	9,980	796	7.98

Tests at all four grade levels were divided into four parts. The first part required the spelling of a dictated list of words. The second part required the composition of a story at the grade 3, 6, and 9 levels, and the planning of an essay at the grade 12 level. The third part involved the producing of sentences from given words or the combining of sentences. The fourth part required the composition of a description or exposition.

All student responses were given in the test booklet. These were hand-scored by classroom teachers.

Step Six: Scoring of the Provincial Sample

A detailed scoring manual was provided to teachers who wished to score their own students' work. Since student responses were actual written productions, rather than selections of correct answers (multiple-choice type), teachers were required to score the tests designated for the provincial sample.

In June teachers of language arts at grade levels 3, 6, 9, and 12 were invited to apply to assist in scoring the provincial sample. In total, 80 teachers were accepted from the 398 that applied. Several teachers at each grade level were selected on the basis of previous experience, either scoring the 1979 writing assessment sample or assisting with the 1982 writing pilot tests. The remaining teachers were selected on a first-applicant basis from major regions in the province.

The scoring of the provincial sample was conducted at the University of Manitoba during July 5 to July 9, 1982. All scorers received training in marking the tests. The scoring was done within grade level groups and according to a set procedure. This was done to minimize scorer variability.

Step Seven: Reporting of Results

The *Preliminary Report* on the Writing Assessment (this report) was produced and distributed to all schools, districts, and divisions.

A *Final Report* will be produced. It will include conclusions and recommendations based on the assessment, and on comparisons with the 1979 assessment.

Step Eight: Interpretation of the Results

Interpretation of the 1982 results will be based in part on 1979 student performance. The 1982 tests included substantial changes from and additions to the 1979 tests. The interpretation of these results will be done with the assistance of experienced language arts teachers and the TAC.

ASSESSMENT IN WRITING

The act of writing is a set of complex cognitive and linguistic activities which cannot be described or measured under any simple taxonomy or by the identification of any particular content; nor is writing the additive accumulation of a set of skills or competencies. In an attempt, however, to assess writing ability, Odell (1981) has defined writing competence as "the ability (1) to discover what one wishes to say and (2) to choose the appropriate language, sentence structures, organizations, and information to achieve a desired purpose with a given audience" (p. 107).

Odell's definition emphasizes the importance of invention, language shaping, and language production in writing. It is significant to note that the definition places the areas of spelling, punctuation and capitalization, and the identification of grammatical constituents as ancilliary to, not central in, writing competence. It seems clear that composing should be the activity central to assessment in writing and that "editing skills" should be assessed only as they support the act of composing. The 1982 Manitoba Writing Assessment, based on this point of view, emphasized the productive and inventive nature of writing rather than the mechanical manipulation of language or of writing conventions.

The definition and emphasis on composing in writing have been derived from research evidence gathered over the past fifty years (see Braddock, Lloyd-Jones, & Schoer, 1962; Mellon, 1978; Odell, 1981). It was decided by the TAC that the provincial assessment in writing should make the description of students' production and discourse skills its primary focus, and should make the assessment of the mechanical skills and conventions of writing secondary. This is the result of the committee's examination of the latest research and practice in the assessment of writing, particulary the work done in the past three years in New Brunswick, Ontario, and British Columbia (see Evans, 1979 and Conry & Jeroski, 1980).

The Validity of Particular Assessment Procedures in Writing

Obviously, the best way to assess a person's writing ability is to observe that person in writing situations and then to judge the effectiveness of the resultant performance. A number of authors in the field of composition have discussed the validity of different types of measures of writing and writing performance and have come to a similar conclusion: The most valid measures of writing ability are samples and descriptions of actual writing (See Cooper & Odell, 1977; Diederich, 1974; Lloyd-Jones, 1978; Straw, 1981). In developing a provincial assessment in writing, our goal was to observe students systematically in writing, to describe their performance, and to judge the appropriateness of the writing product in terms of general development in composition. Implicit in that goal was the assumption that writing performance would be measured in the most valid and meaningful ways possible under the time and economic constraints of a province-wide assessment.

Herein lay the dilemma. Ideally, a description of student writing would include each mental and physical step students undergo while composing: reaction to a stimulus, generation of language, revising the writing, etc. It would also include observation of writing under a number of different situations, such as with different purposes, audiences, and topics. Some researchers in the field have attempted such complete descriptions. Because of the completeness of these descriptions, however, they have been limited to a handful of students; descriptive techniques have not yet been developed that allow such descriptions for large numbers of students as is the case with a province-wide assessment in writing.

On the other hand, large numbers of students can be tested using "objective" examinations (multiple-choice, error correction, usage choice, fill-in, etc.), but these techniques do not describe actual writing performance or ability. Most often, such techniques assess students' conscious knowledge about language structures and rules (rather than how structures are used in writing) and their ability in correcting other people's errors in unfamiliar writing (rather than the ability to avoid error in their own writing). Moreover, these types of tests tell us little about students' composing or language-generating and -shaping ability, particularly beyond the sentence level. Although objective tests are cheap and easy, and easy to analyze, they provide little information about students' composing ability and no description of students' writing performance. Odell (1981) observes that objective tests are "so limited, in some cases so badly conceived, that (they) cannot provide an adequate picture of students' writing performance." He goes on to point out

> that the skills needed to do well on these tests are not the same as the skills needed to do well in writing. For example, multiple choice tests frequently ask students to choose from among several alternatives that someone else has identified. But for writers, the primary problem is not one of choosing from among a set of alternatives but of generating their own alternatives from which they will choose . . . (Objective) tests do not ask students to make such decisions . . . As I have already suggested, this notion of competence is not comprehensive enought to let us distinguish between writers who are competent (even minimally competent) and those who are not. (pp. 107-108)

In a province-wide assessment, it would be impractical to observe the entire writing process in as many different writing situations as would be desirable in order to assemble a complete description of student writing in the province. We must, therefore, sample from the domain of writing skills, choosing the most descriptive, and attempt to assemble our descriptions based on that sample. In order for our descriptions to be as nearly accurate and valid as possible, the tasks used in assessing must be valid, as well as reflective, of writing competence. As pointed out by Cooper (1977), this type of assessment would, whenever possible, be based on students' actual writing and would reflect the choices students make in their own writing. When actual writing samples are impossible to obtain, then tasks should resemble, as closely as possible, those things which writers actually do while composing.

Development of the 1982 Manitoba Writing Assessment Program

The use of objective items on the provincial assessment was limited to circumstances: (1) where it was impossible to gain information from students; actual writing and (2) when tasks that resembled actual composition experiences could not be devised. With the exception of the spelling task, objective items were not used on the 1982 Writing Assessment.

The 1982 Manitoba Writing Assessment Program was based on actual student writing in response to a set of carefully selected stimuli. At three of the four grade levels (3, 6, 9), students were asked to respond in extended writing to two different situations--one designed to elicit a narrative (story), the other designed to elicit a description/exposition (process writing). At the grade twelve level, students were asked to write one descriptive/expository essay and, instead of a second complete writing task, were asked to plan a persuassive essay and write the concluding paragraph. These extended writing tasks were evaluated in a number of ways: by general reaction (General Impression Marking), by assessment of particular qualities of writing (analytic Scales), by the judging of some limited attributes of the writing (Attribute Scales), and by a set of structure and error counts (Descriptive Scales/Error Counts).

The intent of the 1982 Assessment was to measure students' performance in actual writing situation. A number of objectives that appeared on the 1979 Assessment were set aside so that more actual extended writing could take place on the 1982 Assessment. Whenever possible, objectives from the 1979 Assessment were incorporated into the 1982 Assessment; however, the primary goal was to have students engage in actual writing tasks and to base the provincial assessment on that writing. Descriptions of particular techniques for the assessment are given in the scoring manuals and are only briefly described in this report.

PROCEDURES FOR COMPARISON WITH THE 1979 ASSESSMENT

One of the intents of the 1982 Writing Assessment Program is to provide performance comparisons with the 1979 Assessment. Several aspects of the testing procedures, as well as more general problems of measuring performance change over time, complicated the task. The 1982 tests differed from the 1979 tests in terms of some of the objectives assessed and in the actual test questions.

The 1982 tests contained four parts, two required actual written products, one was spelling by dictation, and one involved sentence production. At each grade level comparisons will be made for one of the written products and in spelling. These will be reported in the Final Report.

REFERENCES

Braddock, R., Lloyd-Jones, R. & Schoer, L. Research in written composition. Champaign, Illinois: National Council of Teachers of English, 1962.

Conry, R. & Jeroski, S. The British Columbia assessment of written expression. (General report). The Ministry of Education, British Columbia, 1980.

Cooper, C.R. & Odell, L. (Eds.) Evaluating writing: Describing, measuring, judging. Urbana, Illinois: National Council of Teachers of English, 1977.

Diederich, P.B. Measuring growth in English. Urbana, Illinois: National Council of Teachers of English, 1974.

Evans, P.J.A. Evaluation of writing in Ontario: Grades 8, 12, and 13. Toronto: The Ministry of Education, Ontario, 1979.

Lloyd-Jones, R. Sources of knowledge about literacy. In R. Beach & P.D. Pearson (Eds.), Perspectives on literacy: Proceedings of the 1977 Perspectives on Literacy conference. Minneapolis: College of Education, University of Minnesota, 1978.

Mellon, J.C. A taxonomy of compositional competencies. In R. Beach & P.D. Pearson (Eds.), Perspectives on literacy: Proceedings of the 1977 Perspectives on Literacy conference. Minneapolis: College of Education, University of Minnesota, 1978.

Odell, L. Defining and assessing competence in writing. In C.R. Cooper (Ed.), The nature and measurement of competency in English. Urbana, Illinois: National Council of Teachers of English, 1981.

Straw, S.B. Assessment and evaluation in written composition: A commonsense perspective. In V. Froese & S.B. Straw (Eds.), Research in the language arts: Language and schooling. Baltimore: University Park Press, 1981.

Thomas, V. Teaching spelling: Canadian word lists and instructional techniques (Second Edition). Toronto: Gage, 1979.

CHAPTER 2

GRADE 3 TEST RESULTS

The grade 3 writing test consisted of four parts:

- Part I - Spelling: list of 25 spelling words dictated by the teacher.
- Part II - Story Writing. written product (narration) in response to a picture stimulus.
- Part III - Sentence Writing: sentence production based on lists of words and on short sentences.
- Part IV - Writing Directions: written product (process writing) giving directions on how to do something.

Parts II and IV resulted in a written composition. Both written products were scored according to the same scales and descriptions.

Results of the 1982 assessment are presented in the order specified below. A brief description of the scales and scoring procedures precedes the presentation of results.

WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 8
  General Impression Marking Scale (GIM). 8
  Analytic Scales 8
  Story Writing: GIM results 10
  Story Writing: Analytic Scale Results. 12
  Writing Directions: GIM Results. 16
  Writing Directions: Analytic Scale Results . . . 18

WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 22
  Attribute Marking Scales. 22
  Descriptive Scales. 22
  Error Counts. 23
  Story Writing: Attribute and Descriptive Scales/Error Counts. 24
  Writing Directions: Attribute and Descriptive Scales/Error Counts 28

SPELLING. 32
  Reporting of the 1982 Results. . . . .

SENTENCE WRITING. 34
  Objectives at the Grade 3 Level. 34
  An Example 35
  Results for the Sentence Writing 35
    Section A 37
    Section B 38
    Section C 39
    Section D 40

WRITTEN PRODUCTS:
GENERAL IMPRESSION MARKING AND ANALYTIC SCALES

On the 1982 Manitoba Writing Assessment, students' written products were evaluated in a number of different ways in order to come to conclusions about the writing performance demonstrated by students. All student writing was evaluated on two types of scales: A General Impression Marking Scale (one score) and Analytic Marking Scales (nine scores). Thus, each student paper was assigned ten different scores for each written product (two each in grades 3, 6, and 9; one in grade 12).

The scales are described in detail in the Manual and Scoring Key at each grade level. A summary is given below.

General Impression Marking Scale (GIM)

General Impression Marking (GIM) is a judgement about the overall quality of a student's writing performance, given the purpose of the writing, the audience to which the writing was addressed, and the expectation for that performance at a grade level. On the provincial tests, two purposes were identified for students at grade levels 3, 6, and 9. The first purpose was to relate a story (narration); the second purpose was to describe or explain how to do something (process writing). At the grade 12 level, students were asked only to write to describe or explain a process. At all grade levels and on all tasks, the audience identified for the students was their class.

The GIM is a quick impressionistic score. Each written product from each student was read by two different raters quickly for a general reaction. No marks were placed on the papers themselves. Each rater assigned a score of 1 to 6 (1 - low, 6 - high). If the scores of the two raters did not differ more than one point (e.g., one rater assigned the paper a 5, the other assigned it a 4), the scores were added together (5 + 4 = 9). That summed score is the GIM score reported for students on the writing task. If the scores of the two raters did differ more than one point (e.g., one rater assigned the paper a 2, the other a 5), the paper was read by an arbitrator. (Arbitrators were scorers chosen because of their experience in doing this type of marking.) In the arbitrated cases, the GIM score for a piece of writing was calculated as follows:

$$\frac{\text{FIRST SCORE + SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{GIM SCORE}$$

The lowest score a written product could receive was 2 and the highest was 12.

The GIM score attempts to rate compositions in a general, global manner (See Cooper & Odell, 1977 for a complete discussion). Covered by the GIM are any qualities that impress the rater including the quality of thought evidenced in the paper, the effectiveness of the overall organization, the effectiveness of the overall presentation, and the effectiveness of the style and word choice.

Analytic Scales

Like general impression marking, the Analytic Scales are quick, qualitative judgements about the effectiveness of a piece of writing. Analytic Scales, however, are measures of the effectiveness of particular qualities of a paper. For the provincial writing assessment, nine different qualities were

selected for testing, five rhetorical or discourse categories, four mechanics and convention categories. These scales were developed because they seemed to be the qualities most often evaluated in writing by teachers and because they were objectives most often identified as being taught by teachers.

Like the GIM, a paper was assigned a score between 1 and 6 (1 - low, 6 - high), but for each of the nine qualities.

Rhetorical Categories

1. Organization and Sequence
2. Wording and Style
3. Focus or Theme
4. Ideas
5. Choice of Details and Elaboration of Ideas

Mechanics and Conventions Categories

6. Sentence Structure
7. Usage
8. Punctuation and Capitalization
9. Spelling

A scorer, usually different from the GIM scorers, read each written product for the Analytic Scales. Specific descriptions of the qualities associated with high, middle, and low papers are given in each grade level Manual and Scoring Key, as well as examples of papers. (For a more complete discussion of analytic scales in general, refer to Diederich, 1974, and Cooper & Odell, 1977.)

For the provincial sample, every paper was rated on all nine analytic scales by at least one rater. A subset of papers at each grade level was marked analytically by a second rater in order to calculate the amount of agreement or reliability between independent raters.

These two scoring procedures (GIM and Analytic) yielded ten different scores for each written product. At the grade 3, 6, and 9 levels, students wrote two pieces of extended writing so that each student in those grades received 20 different scores (10 for each writing product). At the grade 12 level, the students were asked to write one piece of extended writing so that each student received 10 different scores. This is summarized in Table 2 (p. 10).

The GIM and Analytic score results at the grade 3 level are based on scoring the total provincial sample of 1333 test papers.

TABLE 2

GIM AND ANALYTIC SCORING OF WRITTEN PRODUCTS

Scales	Narrative Writing 3	6	9	12	Descriptive/Expository Writing 3	6	9	12
GIM (Scores range from 2 to 12)	x	x	x	-	x	x	x	x
Analytic (Scores range from 1 to 6)								
1. Organization and sequence	x	x	x	-	x	x	x	x
2. Wording and style (Word Choice)	x	x	x	-	x	x	x	x
3. Focus or theme	x	x	x	-	x	x	x	x
4. Ideas	x	x	x	-	x	x	x	x
5. Choice of Details & Elaboration of Ideas	x	x	x	-	x	x	x	x
6. Sentence Structure	x	x	x	-	x	x	x	x
7. Usage	x	x	x	-	x	x	x	x
8. Punctuation & Capitalization	x	x	x	-	x	x	x	x
9. Spelling	x	x	x	-	x	x	x	x

Story Writing: GIM Results

Students at the grade 3 level were asked to write a story using the picture provided as a stimulus (Part II of the test). Teachers were instructed to read the directions aloud to their students. Students were given approximately 35 minutes to write their stories.

The results of the GIM scoring for the story are presented on page 11. Scores for the two scorers combined ranged from 2 to 12, with scores, 2, 3, and 4 described as "Low"; 5, 6, 7, 8, and 9 described as "Middle"; and 10, 11, and 12 described as "High".

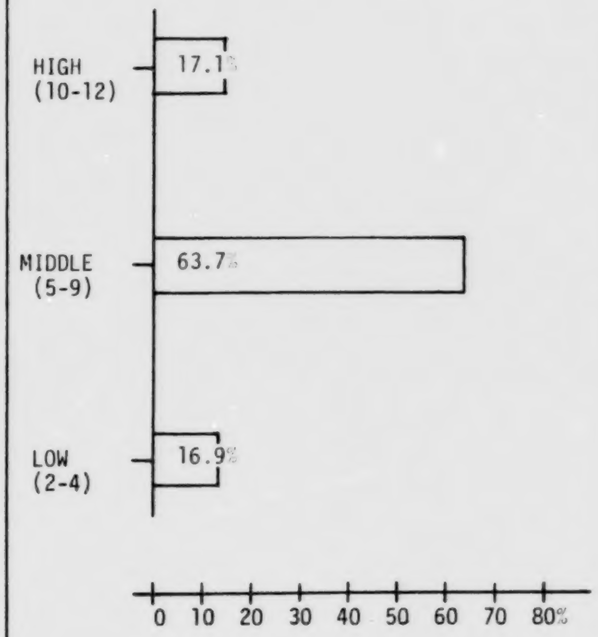
The results are presented in three ways:

1. The percentage of students obtaining each GIM score: e.g., 2.6% of students received a combined GIM score of 12 (both scorers awarded a 6 to these stories).
2. The percentage of students obtaining GIM scores categorized as High, Middle, and Low: e.g., 17.1% of students received combined GIM scores of 10, 11, or 12 (High).
3. The average (mean) performance of students on the GIM scale, (and the standard deviation). These are based on the scale of 2 to 12.

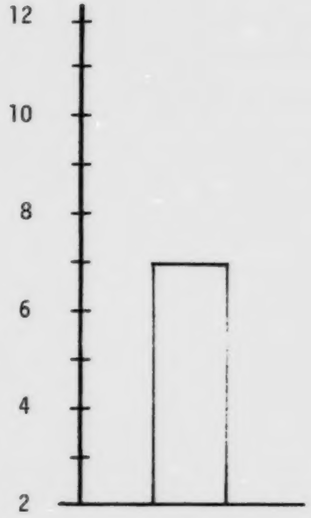
STORY WRITING: 1982 GIM RESULTS (Grade 3)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
HIGH	12	2.6
	11	7.3
	10	7.2
MIDDLE	9	12.6
	8	9.0
	7	17.7
	6	11.9
	5	12.5
LOW	4	8.4
	3	4.8
	2	3.7
Insufficient Information		2.5

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 6.99

Standard Deviation = 2.51

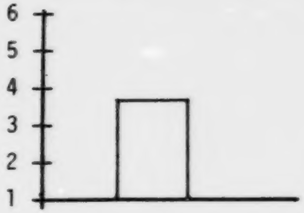
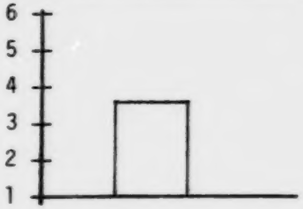
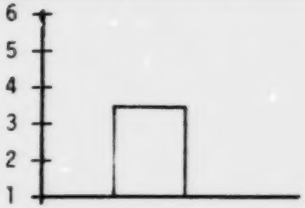
Story Writing: Analytic Scale Results

The results of the analytic scoring for the grade 3 story are presented on pages 13 to 15. All of the scores range on a scale from 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High".

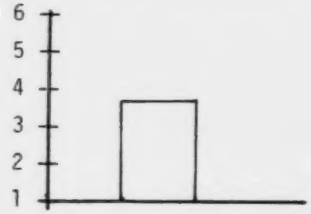
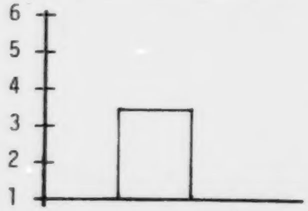
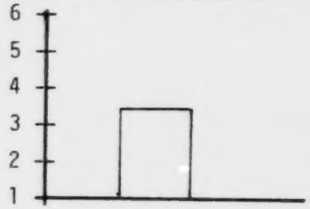
Results are presented in two ways:

1. Percentage of students obtaining each score: e.g., 8.0% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students, and standard deviation, based on the scale of 1 to 6: e.g., the average of students was 3.73 (standard deviation was 1.29) on Organization and Sequence.

STORY WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 3)

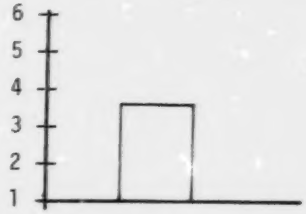
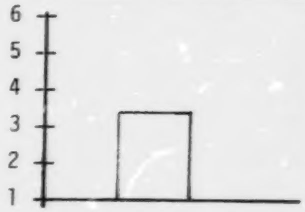
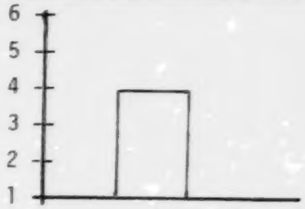
Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	8.0%	 Average = 3.73 Standard Deviation = 1.29
		5	22.5	
	MIDDLE	4	25.6	
		3	24.6	
	LOW	2	14.9	
		1	3.6	
	Insufficient Information		0.9	
2. Word Choice	HIGH	6	7.8%	 Average = 3.66 Standard Deviation = 1.20
		5	16.3	
	MIDDLE	4	27.2	
		3	32.9	
	LOW	2	12.0	
		1	2.9	
	Insufficient Information		0.8	
3. Focus or Theme	HIGH	6	6.2%	 Average = 3.57 Standard Deviation = 1.34
		5	22.1	
	MIDDLE	4	22.8	
		3	25.3	
	LOW	2	15.6	
		1	7.0	
	Insufficient Information		1.1	

STORY WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 3)

Analytic Marking Scale: Objective	Analytic Marking Scale: Scale		Percentage of Students Obtaining Each Score	Average Performance
4. Ideas	HIGH	6	8.6	
		5	17.7	
	MIDDLE	4	21.7	
		3	29.8	
	LOW	2	16.1	
		1	4.6	
	Insufficient Information		1.6	Average = 3.59 Standard Deviation = 1.32
5. Choice of Details and Elaboration of Ideas	HIGH	6	7.0	
		5	15.2	
	MIDDLE	4	19.8	
		3	30.7	
	LOW	2	18.8	
		1	7.0	
	Insufficient Information		1.6	Average = 3.39 Standard Deviation = 1.34
6. Sentence Structure	HIGH	6	8.2	
		5	17.4	
	MIDDLE	4	21.1	
		3	24.1	
	LOW	2	18.3	
		1	10.2	
	Insufficient Information		0.9	Average = 3.42 Standard Deviation = 1.44

STORY WRITING: 1982 ANALYTIC SCALE RESULTS

(Grade 3)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	8.6	
		5	21.2	
	MIDDLE	4	23.5	
		3	20.5	
	LOW	2	16.7	
		1	7.7	
	Insufficient Information		1.9	Average = 3.61 Standard Deviation = 1.42
8. Punctuation and Capitalization	HIGH	6	8.6	
		5	19.8	
	MIDDLE	4	19.6	
		3	18.0	
	LOW	2	19.8	
		1	13.2	
	Insufficient Information		0.8	Average = 3.39 Standard Deviation = 1.54
9. Spelling	HIGH	6	14.3	
		5	24.1	
	MIDDLE	4	22.6	
		3	18.4	
	LOW	2	12.2	
		1	5.9	
	Insufficient Information		2.5	Average = 3.92 Standard Deviation = 1.44

Writing Directions: GIM Results

Students at the grade 3 level were asked to write directions about how to do something, using as suggested topics: How to make a snowman, How to care for my pet, or How to make my favorite food (Part IV of the test). Teachers were instructed to read the directions aloud to their students. Students were given approximately 25 minutes to write their directions.

These student compositions were scored using the same GIM and analytic scales as were used for the stories (see pages 8-10).

The results of the GIM scoring are presented on page 17. Scores for the two scorers combined ranged from 2 to 12, with scores 2, 3, and 4 categorized as "Low"; 5, 6, 7, 8, and 9 as "Middle"; and 10, 11, and 12 as "High".

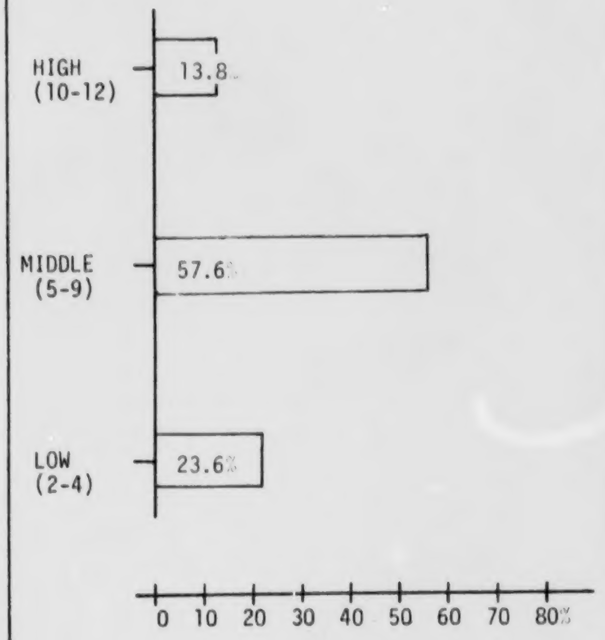
The results are depicted in three ways:

1. The percentage of students obtaining each GIM score: e.g., 2.2% of students received a combined GIM score of 12.
2. The percentage of students obtaining GIM scores categorized as High, Medium, and Low: e.g., 13.8% of students received scores in the High category (GIM score of 10, 11, or 12).
3. The average (mean) performance of students on the GIM scale (and the standard deviation).

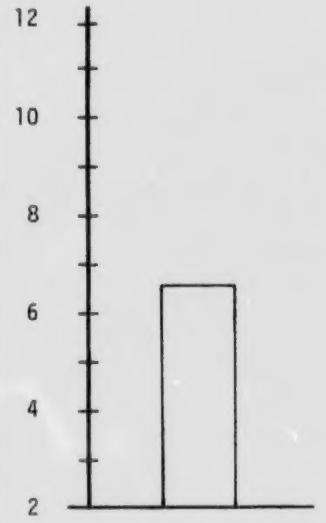
WRITING DIRECTIONS: 1982 GIM RESULTS (Grade 3)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
HIGH	12	2.2%
	11	5.8
	10	5.8
MIDDLE	9	10.6
	8	9.2
	7	13.7
	6	11.7
	5	12.4
LOW	4	10.0
	3	8.2
	2	5.4
Insufficient Information		5.1

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 6.54

Standard Deviation = 2.62

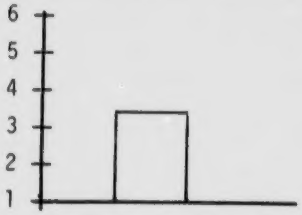
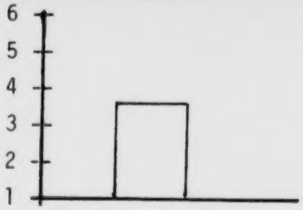
Writing Directions: Analytic Scale Results

The results of the analytic scoring for the grade 3 process writing are presented on pages 19 to 21. All of the analytic scores are on a scale of 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High". These scales are identical to those used in evaluating the narrative writing (see pages 8 - 10 for a detailed description).

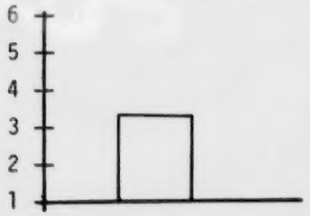
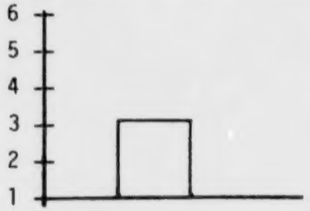
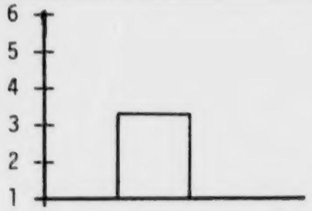
Results for each analytic scale are depicted in two ways:

1. Percentage of students obtaining each score: e.g., 4.8 of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (Mean) performance of students on each analytic scale (and standard deviation) based on the scale from 1 to 6: e.g., the average performance was 3.49 (standard deviation was 1.26) on Organization and Sequence.

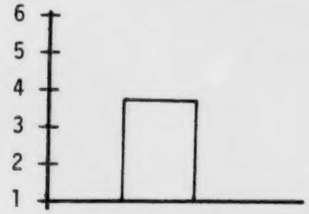
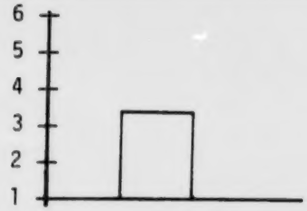
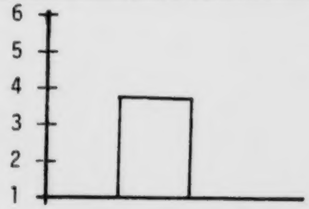
WRITING DIRECTIONS: 1982 ANALYTIC SCALE RESULTS (Grade 3)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	4.8	
		5	17.8	
	MIDDLE	4	24.2	
		3	28.0	
	LOW	2	16.8	
		1	5.0	
	Insufficient Information		3.4	Average = 3.49 Standard Deviation = 1.26
2. Word Choice	HIGH	6	4.4	
		5	12.6	
	MIDDLE	4	28.5	
		3	35.1	
	LOW	2	13.5	
		1	3.2	
	Insufficient Information		2.9	Average = 3.48 Standard Deviation = 1.12
3. Focus or Theme	HIGH	6	5.5	
		5	23.5	
	MIDDLE	4	24.8	
		3	25.6	
	LOW	2	11.4	
		1	5.5	
	Insufficient Information		3.7	Average = 3.69 Standard Deviation = 1.28

WRITING DIRECTIONS: 1982 ANALYTIC SCALE RESULTS (Grade 3)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	4.7%	 Average = 3.33 Standard Deviation = 1.24
		5	12.3	
	MIDDLE	4	24.5	
		3	31.3	
	LOW	2	17.1	
		1	6.8	
	Insufficient Information		3.4	
5. Choice of Details and Elaboration of Ideas		6	4.7%	 Average = 3.03 Standard Deviation = 1.34
	HIGH	5	9.0	
		4	19.4	
	MIDDLE	3	26.8	
	LOW	2	24.0	
		1	12.4	
	Insufficient Information		3.7	
6. Sentence Structure	HIGH	6	2.8%	 Average = 3.17 Standard Deviation = 1.32
		5	15.6	
	MIDDLE	4	19.3	
		3	28.2	
	LOW	2	19.8	
		1	11.3	
	Insufficient Information		3.1	

WRITING DIRECTIONS: 1982 ANALYTIC SCALE RESULTS (Grade 3)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	4.1%	
		5	25.0	
	MIDDLE	4	23.9	
		3	27.2	
	LOW	2	11.4	
		1	5.0	
	Insufficient Information		3.3	Average = 3.67 Standard Deviation = 1.24
8. Punctuation and Capitalization	HIGH	6	5.7%	
		5	18.4	
	MIDDLE	4	20.2	
		3	21.9	
	LOW	2	15.6	
		1	14.7	
	Insufficient Information		3.5	Average = 3.30 Standard Deviation = 1.48
9. Spelling	HIGH	6	7.6%	
		5	24.2	
	MIDDLE	4	26.3	
		3	23.5	
	LOW	2	10.3	
		1	4.7	
	Insufficient Information		3.4	Average = 3.81 Standard Deviation = 1.28

WRITTEN PRODUCTS:

ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS

In addition to the qualitative judgements made on each paper using the GIM and Analytic Scales, a randomly selected sub-sample of papers was examined for more intensive structural, linguistic, attribute, and error descriptions. These in-depth analyses were carried out by the same group of raters employed for the other marking tasks. A total of 70 student papers were selected at the grade 3 level for this intensive scoring.

The rationale for these types of analyses was to make assessments regarding the kinds of structures students used in their actual writing and the kinds of errors most frequently made by students in actual writing situations. Also, they were conducted in order to draw some conclusions about the number of students demonstrating serious problems in the structural aspects of their writing. The choice of structures, errors, and attributes was based on the survey of writing done by Hunt (See Cooper & Odell, 1977), on the work of Evans (1979) in the Ontario assessment, and on the objectives identified by practicing teachers as the ones of greatest interest and use to them in their classrooms.

Attribute Marking Scales

Seven different attribute marking scales were identified as being of interest to teachers. The scales themselves were evaluations about whether or not a student was experiencing serious (in need of remedial help) difficulty in particular areas. For each paper evaluated, the rater was asked to rate the composition on the seven attributes, each on a three-point scale:

3 - No errors;

2 - An occasional error, but not rated as a serious problem;

1 - Serious problem (Student is in need of special remedial help);

X - Insufficient information (Structures not used or insufficient writing for an evaluation to be made).

Five of the seven areas were evaluated at all grade levels. The last two (use of contractions and use of quotation marks in direct discourse) were evaluated only at grade levels 3 and 6. The seven attributes were:

1. Run-on sentences
2. Use of complete sentences
3. Paragraphing
4. Internal punctuation
5. Capitalization
6. Use of contractions
7. Use of quotation marks in direct discourse

Descriptive Scales

The descriptive scales were counts of particular structures used by students. The following counts were made for each student composition in grades 3, 6, and 9. Grade 12 compositions were tallied on all but the last one.

1. Total number of words written
2. Total number of T-units (independent clauses) written
3. Total number of dependent clauses written
4. Total number of adverbial clauses written

5. Total number of relative clauses written
6. Total number of noun clauses written
7. Total number of modifiers (not counted due to the large number)
8. Total number of instances of direct discourse

The particular counts were chosen because they seem closely related both to composition quality and to syntactic fluency (See Hunt in Cooper & Odell, 1977). The purpose was to judge whether or not students easily and fluently used these structures in their actual writing.

From these counts, a number of transformed scores were calculated to produce twelve indices of writing maturity in students. In general, the more fluent, mature writers will demonstrate higher scores on each of these indices, whereas weaker, less mature writers will score lower. These scores should be viewed, however, only as indicators of how students handle the syntax and grammar of the language in free writing situations. The question answered by these data is: Do students use these particular structures in their own writing? The question of whether or not students can use these structures is left to be answered by the sentence writing tasks (Part III).

The scores calculated included:

Words per T-unit
Words per clause
Clauses per T-unit
Dependent clauses per T-unit
Adverbial clauses per T-unit
Adverbial clauses per 100 words
Relative clauses per T-unit
Relative clauses per 100 words
Noun clauses per T-unit
Noun clauses per 100 words
Direct discourse per T-unit
Direct discourse per 100 words

Error Counts

Error counts, though tedious to do, can give information about how well students avoid error in actual writing situations. These counts are realistic alternatives to the use of objective (particularly multiple-choice type) items in evaluating writing. Error counts can give greater information about the choices that students are making in actual writing tasks.

Each of the error categories are calculated in these ways:

average number of errors per composition,
average number of errors per T-unit (i.e. per independent clause), and
average number of errors per 100 running words

The error counts obtained were:

1. Number of sentence fragments
2. Number of incorrectly placed modifiers
3. Number of sentences not beginning with a capital letter
4. Number of proper nouns not capitalized
5. Number of sentences not ending with appropriate punctuation
6. Number of internal comma errors
7. Number of faulty subject-verb agreement
8. Number of verb tense errors
9. Number of word order errors

Similarly to the descriptive scales a number of scores were calculated from the error counts. But, opposite to the descriptive scales, higher scores on the scale reflect greater error rates.

The calculated scores were:

Sentence fragments per T-unit
Sentence fragments per 100 words
Incorrectly placed modifiers per T-unit
Incorrectly placed modifiers per 100 words
Beginning capitalization errors per T-unit
Beginning capitalization errors per 100 words
Proper nouns not capitalized per T-unit
Proper nouns not capitalized per 100 words
End punctuation errors per T-unit
End punctuation errors per 100 words
Comma errors per T-unit
Comma errors per 100 words
Faulty agreement per T-unit
Faulty agreement per 100 words
Verb tense errors per T-unit
Verb tense errors per 100 words
Word order errors per T-unit
Word order errors per 100 words

Story Writing: Attribute and Descriptive Scales/Error Counts

The results for the grade 3 stories (Part II) are presented for:

The 7 attribute scales (p.25),

The 8 descriptive scales (p. 26),

The 12 indices calculated from the descriptive scales (p.26),

The 9 error counts (p.27), and

The 18 indices calculated from the error counts (p.27).

Since there are a large number of scales, the results have been tabulated in a compacted form in the following tables.

The average (mean) and standard deviation are reported for each scale and index. For the attribute scales, the percentage of students obtaining each score is reported: e.g., 27.1% of students obtained a score of 3 (no errors) on the attribute scale "Run-on sentences", and the average is 2.16 on the scale of 1 to 3 (standard deviation is 0.61).

STORY WRITING: 1982 ATTRIBUTE SCALES RESULTS (Grade 3.

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
1. Run-on sentences				
No errors	3	27.1		
An occasional error	2	61.4	2.16	0.61
Serious problem	1	11.4		
Insufficient information		0.0		
2. Use of complete sentences				
No errors	3	32.9		
An occasional error	2	60.0	2.26	0.58
Serious problem	1	7.1		
Insufficient information		0.0		
3. Paragraphing				
No errors	3	12.9		
An occasional error	2	50.0	1.96	0.61
Serious problem	1	15.7		
Insufficient information		21.4		
4. Internal punctuation				
No errors	3	11.4		
An occasional error	2	60.0	1.91	0.58
Serious problem	1	20.0		
Insufficient information		8.6		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
5. Capitalization				
No errors	3	24.3		
An occasional error	2	61.4	2.12	0.61
Serious problem	1	12.9		
Insufficient information		1.4		
6. Use of contractions*				
No errors	3	27.1		
An occasional error	2	14.3	2.66	0.48
Serious problem	1	0.0		
Insufficient information		58.6		
7. Use of quotation marks in direct discourse*				
No errors	3	12.9		
An occasional error	2	31.4	2.03	0.67
Serious problem	1	11.4		
Insufficient information		44.3		

*Many students did not use these particular features in their writing; this was categorized as "Insufficient information".

STORY WRITING: 1982 DESCRIPTIVE SCALE RESULTS (Grade 3)

Descriptive Scale (Number of)	Average Number Mean	Standard Deviation
1. Words	139.36	71.68
2. T-units	17.47	8.94
3. Dependent Clauses	2.76	2.55
4. Adverbial clauses	1.74	2.00
5. Relative clauses	0.21	0.54
6. Noun clauses	0.86	1.55
7. Modifiers	(not counted)	
8. Instances of direct discourse	1.44	2.15

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	8.22	1.93
Words per clause*	7.18	1.62
Clauses per T-unit	1.15	.13
Dependent clauses per T-unit	0.15	0.13
Adverbial clauses per T-unit	0.089	0.088
Adverbial clauses per 100 words	1.08	1.01
Relative clauses per T-unit	0.016	0.046
Relative clauses per 100 words	0.19	0.52
Noun clauses per T-unit	0.041	0.074
Noun clauses per 100 words	0.51	1.038
Direct discourse per T-unit	0.085	0.130
Direct discourse per 100 words	1.12	1.70

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

STORY WRITING: 1982 ERROR COUNTS (Grade 3)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.055	0.159
Sentence fragments per 100 words	0.58	1.59
Incorrectly placed modifiers per T-unit	.006	0.028
Incorrectly placed modifiers per 100 words	0.071	0.333
Beginning capitalization errors per T-unit	0.25	0.26
Beginning capitalization errors per 100 words	3.13	3.33
Proper nouns not capitalized per T-unit	0.051	0.136
Proper nouns not capitalized per 100 words	0.69	1.83
End punctuation errors per T-unit	0.25	0.27
End punctuation errors per 100 words	3.17	3.23
Comma errors per T-unit	0.18	0.21
Comma errors per 100 words	2.07	2.47
Faulty agreement per T-unit	0.015	0.055
Faulty agreement per 100 words	0.17	0.59
Verb tense errors per T-unit	0.027	0.068
Verb tense errors per 100 words	0.35	0.90
Word order errors per T-unit	0.030	0.115
Word order errors per 100 words	0.34	1.23

Error Counts (Number of)		Percentage of Students Obtaining each Count	Average Error Count Mean	Average Error Count Standard Deviation
1. Sentence fragments	0	70.0%	0.56	1.00
	1	14.3		
	2 or more	15.6		
2. Incorrectly placed modifiers	0	94.3	0.09	0.37
	1	2.9		
	2 or more	2.9		
3. Sentences not beginning with capital letter	0	25.7	4.00	4.58
	1	11.4		
	2 or more	62.9		
4. Proper nouns not capitalized	0	81.4	0.94	2.52
	1	2.9		
	2 or more	15.7		
5. Sentences not ending with appropriate punctuation	0	20.0	4.30	5.07
	1	10.0		
	2 or more	70.0		
6. Internal comma errors	0	31.4	2.70	3.63
	1	11.4		
	2 or more	57.2		
7. Instances of faulty subject/verb agreement	0	90.0	0.19	0.73
	1	7.1		
	2 or more	2.8		
8. Verb tense errors	0	74.3	0.37	0.80
	1	18.6		
	2 or more	7.1		
9. Word order errors	0	84.3	0.31	0.89
	1	7.1		
	2 or more	8.5		

Writing Directions: Attribute and Descriptive Scales/Error Counts

The results for the process writing part of the grade 3 test (Part IV) are presented for the scales and calculated indices. The scales are identical to those used for the stories (for a description see pages 22 to 24):

The 7 attribute scales (p.29),
The 8 descriptive scales (p.30),
The 12 indices calculated from the descriptive scales (p.30),
The 9 error counts (p.31), and
The 18 indices calculated from the error counts (p.31).

Due to the large number of scales and indices the results have been tabulated in a compacted form in the following tables. The average (mean) and standard deviation are reported for each scale and index. As well, for the attribute scales, the percentage of students obtaining each score is reported: e.g., 38.6% of students obtained a score of 3 (no errors) on the attribute scale "Run-on Sentences", and the average is 2.27 on the scale of 1 to 3 (standard deviation 0.68).

WRITING DIRECTIONS: 1982 ATTRIBUTE SCALES RESULTS (Grade 3)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
1. Run-on sentences				
No errors	3	38.6%		
An occasional error	2	45.7	2.27	0.68
Serious problem	1	12.9		
Insufficient information		2.9		
2. Use of complete sentences				
No errors	3	30.0		
An occasional error	2	45.7	2.09	0.73
Serious problem	1	21.4		
Insufficient information		2.9		
3. Paragraphing*				
No errors	3	15.7		
An occasional error	2	32.9	2.10	0.66
Serious problem	1	10.0		
Insufficient information		41.5		
4. Internal punctuation				
No errors	3	18.6		
An occasional error	2	48.6	2.02	0.66
Serious problem	1	17.1		
Insufficient information		15.8		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
5. Capitalization				
No errors	3	27.1%		
An occasional error	2	45.7	2.09	0.71
Serious problem	1	18.6		
Insufficient information		8.6		
6. Use of contractions*				
No errors	3	11.4		
An occasional error	2	7.1	2.40	0.74
Serious problem	1	2.9		
Insufficient information		78.6		
7. Use of quotation marks in direct discourse*				
No errors	3	-		
An occasional error	2	-	-	-
Serious problem	1	-		
Insufficient information		95.8		

*Many students did not use these particular features in their writing; this was categorized as "Insufficient information".

WRITING DIRECTIONS: 1982 DESCRIPTIVE SCALE RESULTS (Grade 3)

	Descriptive Scale (Actual counts)	Average Number	
		Mean	Standard Deviation
1.	Words	73.31	38.96
2.	T-units	7.76	4.34
3.	Dependent Clauses	1.66	2.40
4.	Adverbial clauses	1.41	2.10
5.	Relative clauses	0.01	0.12
6.	Noun clauses	0.14	0.64
7.	Modifiers	(not counted)	
8.	Instances of direct discourse	0.0	0.0

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	10.30	4.56
Words per clause*	8.89	4.47
Clauses per T-unit	1.18	0.30
Dependent clauses per T-unit	0.19	0.26
Adverbial clauses per T-unit	0.16	0.22
Adverbial clauses per 100 words	1.49	1.88
Relative clauses per T-unit	0.002	0.015
Relative clauses per 100 words	0.012	0.103
Noun clauses per T-unit	0.018	0.069
Noun clauses per 100 words	0.15	0.62
Direct discourse per T-unit	0.0	0.0
Direct discourse per 100 words	0.0	0.0

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

WRITING DIRECTIONS: 1982 ERROR COUNTS (Grade 3)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.23	0.94
Sentence fragments per 100 words	1.31	3.96
Incorrectly placed modifiers per T-unit	0.0	0.0
Incorrectly placed modifiers per 100 words	0.0	0.0
Beginning capitalization errors per T-unit	0.38	0.40
Beginning capitalization errors per 100 words	3.87	4.09
Proper nouns not capitalized per T-unit	0.010	0.045
Proper nouns not capitalized per 100 words	0.10	0.45
End punctuation errors per T-unit	0.32	0.37
End punctuation errors per 100 words	3.69	4.63
Comma errors per T-unit	0.15	0.22
Comma errors per 100 words	1.64	2.39
Faulty agreement per T-unit	0.005	0.021
Faulty agreement per 100 words	0.053	0.216
Verb tense errors per T-unit	0.014	0.083
Verb tense errors per 100 words	0.21	1.35
Word order errors per T-unit	0.024	0.128
Word order errors per 100 words	0.28	1.56

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Average Error Count Standard Deviation
1. Sentence fragments	0	70.0%	0.57	1.21
	1	20.0		
	2 or more	10.0		
2. Incorrectly placed modifiers	0	100.0	0.00	0.00
	1	0.0		
	2 or more	0.0		
3. Sentences not beginning with Capital letters	0	28.6	2.26	2.48
	1	20.0		
	2 or more	51.4		
4. Proper nouns not capitalized	0	94.3	0.09	0.37
	1	2.9		
	2 or more	2.9		
5. Sentences not ending with appropriate punctuation	0	27.1	2.23	2.87
	1	17.1		
	2 or more	45.8		
6. Internal comma errors	0	52.9	1.16	1.59
	1	14.3		
	2 or more	32.8		
7. Instances of faulty subject/verb agreement	0	94.3	0.06	0.23
	1	5.7		
	2 or more	0.0		
8. Verb tense errors	0	94.3	0.07	0.31
	1	4.3		
	2 or more	1.4		
9. Word order errors	0	92.9	0.14	0.62
	1	4.3		
	2 or more	2.8		

SPELLING

The purpose of the spelling assessment on the 1982 Writing Assessment Program was to determine if students could spell correctly the words they actually used in their writing as well as the words they were likely to need in their writing. Spelling in the actual written products was assessed by one of the analytic scales (Analytic Scales, p.8). In order to assess the performance of students on the spelling of words they were likely to need in their general writing (but may not have used in the writing tasks provided on the test), words were randomly selected from Ves Thomas' Canadian Word Lists (1979). This particular set of word lists assigns words to different grade levels based on the frequency of the appearance of these words in the writing of students at a particular grade level. For example, the grade 3 list is composed of words that a high number of Canadian grade 3 students use in their own writing.

In addition to the words selected from the Thomas Canadian Word Lists, a number of words were selected from the 1979 Manitoba Writing Assessment Program so that comparisons between student performance in 1979 and 1982 could be made. These 1979 words were originally selected not because they should be in students' *writing* vocabularies but because they conformed to particular patterns and appeared in grade level *reading* materials: that is, they were words that students could be expected to recognize and pronounce (but not necessarily be able to use or spell in writing). The rationale for the selection of the 1982 spelling words seemed stronger in that it reflected words that students were likely to use in composing rather than in reading.

The spelling test used in the 1982 assessment contained two separate tests (Part I of the test). One set of items (those marked with a * in the *Manual and Scoring* Key) was taken from the 1979 Manitoba Writing Assessment. The total number of these words is 7.

The balance of the words were randomly selected from the Thomas list at the grade 3 level. There were 18 words obtained in this way for the 1982 component.

Two scores were given on the test, one for the items from 1979 (Total possible = 7), another for the items selected for 1982 (Total possible = 18). Total overall performance (Total possible = 25) was determined by adding the two scores together. The results are based on the provincial sample of 1333 grade 3 students.

Reporting of the 1982 Results

The average (mean) score is reported for the scores based on the 7 items, on the 18 items, and on the 25 items (combination of 7 and 18 items). A frequency distribution indicating the percentage of students obtaining each score is presented (e.g., 18.8% of students spelled 22 or 23 of the words correctly). Finally, the percentage of students spelling each word correctly is indicated on page 33 (e.g., 79% of students spelled "horses" correctly.

SUMMARY OF THE 1982 SPELLING RESULTS

(Grade 3)

Score Correct out of Total Number of Words (25)	Percentage of Students Obtaining Each Score
24-25	42.5%
22-23	18.8
20-21	11.7
18-19	7.9
16-17	5.1
14-15	4.2
12-13	2.8
10-11	1.9
8-9	1.7
6-7	1.1
4-5	0.8
2-3	0.8
0-1	0.9

Average Score Correct	20.8
Standard Deviation	5.22
Average Percentage Correct	83.2%

Words Selected From the 1979 Assessment		Words Selected For the 1982 Assessment	
Items	Percentage of 1982 students obtaining item correct	Items	Percentage of 1982 students obtaining item correct
8. it's	88.7%	1. horses	79.0%
10. rabbits	75.4	2. children	81.0
12. I'll	81.7	3. snake	87.0
14. because	79.0	4. girls	83.1
17. shelter	69.0	5. boys	93.3
20. catch	77.3	6. weeks	89.8
23. swimming	72.9	7. real	86.8
		9. landed	89.8
		11. wouldn't	65.3
		13. lights	85.5
		15. window	91.1
		16. wanted	81.9
		18. hit	94.6
		19. planet	71.0
		21. tell	93.5
		22. off	94.0
		24. little	87.1
		25. only	82.6

Average Correct	5.4	Average Correct	15.4
Standard Deviation	1.81	Standard Deviation	3.64
Average Percentage Correct	77.7%	Average Percentage Correct	85.3%

SENTENCE WRITING

The Technical Advisory Committee developed a set of sentence writing tasks to assess performance on selected structural, syntactic, usage, and, at some levels, mechanics objectives. The particular format of the sentence writing tasks was chosen for five reasons:

- They closely resembled the kind of thinking and shaping a writer probably goes through when composing.
- They allowed for a variety of answers which could be labeled "correct" so that a student was not limited to a single correct answer and could create the most effective sentence(s) for the situation and information provided.
- The structurally correct answers in most cases could be generally categorized from least mature (or fluent) to most mature (or fluent).
- The open format centred on what students could do in particular circumstances rather than on what they could not do.
- The format could allow for the assessment of whether students could use particular structures.

Performance on the sentence writing tasks could be compared with the descriptive scales in the actual writing tasks in order to determine if students in fact did use the targeted structures in free writing.

Although a number of objectives (or targeted structures) were common across all four grade levels, the particular structures or conventions assessed at each grade level were based on objectives that practicing teachers felt were applicable at that level.

Objectives at the Grade 3 Level

Objectives at the Grade 3 level are listed below; the sections comprised Part III of the student test.

Section A - Objective 1: Writing complete sentences
Section A - Objective 2: Capitalization and punctuation of sentences
Section B - Objective 1: Writing questions
Section B - Objective 2: Capitalization and punctuation of questions
Section C - Objective 1: Subject coordination
Section C - Objective 2: Using commas in a series
Section D - Objective 1: Using relativization
Section D - Objective 2: Using adverbial subordination

Several sentences were required for each section but some sentences were scored for two objectives.

The objectives fall into three basic areas: sentence sense (A1, B1), capitalization and punctuation (A2, B2, C2), and syntactic maturity (C1, D1, D2). All students should have some understanding of these areas.

Students' sentences were scored according to several categories; these were denoted a, b, c, etc. It is likely that the most careful, mature writers would respond consistently with answers that would fall into category "a" of each objective. Less mature writers might have a profile of mixed performance (e.g., demonstrate sentence sense and knowledge of punctuation, but not consistently demonstrate the ability to capitalize correctly or use the relatively mature structures such as adverbial subordination). The fact that a student did not consistently answer with the most mature structures does not necessarily mean that she/he does not know how to employ the structure or convention, rather that the structure or convention is not readily accessible in the written mode. The very weakest students might demonstrate little fluency in the structures they write or choose. The tasks presented in the assessment allow for divergent answers so that the teacher can assess the level of maturity or fluency of the student(s).

An Example

For Item 1 of Section C of the Grade 3 test (Objective C1), the student was asked to combine the following two sentences using the word "and".

1. The books were lying on the table.
   The pencils were lying on the table.

The most fluent writers probably responded with answers such as:

> The books and the pencils were lying on the table.

This answer was categorized a: Combined using subject coordination.

Somewhat less fluent writers might have responded with an answer such as:

> The books were lying on the table and the pencils were lying on the table.

This answer would be categorized b: Combined using clausal coordination. The answer is still structurally correct and grammatical; it is not as efficient or fluent as the first answer. The categorization scheme for this objective allows the teacher to judge a structure as "correct", but "less mature".

Other writers might have answered:

> Not only were the books lying on the table, but the pencils were also.

Here, the writer has clearly not followed directions--but the purpose of the writing assessment was not to evaluate direction following behavior, but writing behavior. The student has written a structurally correct, grammatical, and meaningful sentence, and should be given credit for it. But, it does not fall into either of the previous categories. This answer would be categorized as c: Combined in some other fashion. The maturity or fluency level here is not apparent from the category description, but the scheme does give the writer credit for meaningful writing.

The final category, d: Incorrectly combined, was only used if the student wrote an answer that did not make sense or was not grammatical.

The categories for scoring varied according to the particular objective, but in all cases the scheme was similar to the example presented.

The complete category descriptions and examples of answers are given in the grade 3 Manual and Scoring Key.

Results for the Sentence Writing

At the grade 3 level students were asked to do two sentence writing tasks (items) for each of the objectives. The sentences in sections A and B were also scored for a second objective. In both cases the second objective was regarding correct capitalization and punctuation.

The results are presented for each sentence and for each objective. Since the scoring categories for each objective were not based on a numerical scale averages cannot be used.

Instead reported are the percentages of students whose sentences were categorized in each of the scoring categories (the first scoring category is always the targeted structure of the objective). Also reported is the average of the percentages for a particular category for the two items pertaining to an objective. For example, in Section A 96.2% of student sentences were categorized as "a-Complete sentence" for Item 1 and 94.8% for Item 2. On average, 95.5% of student sentences were categorized as "a-Complete sentence" for the two sentences combined and thus for the objective: Writing complete sentences.

Results are based on scoring all 1333 papers in the provincial sample at the grade 3 level.

SENTENCE WRITING: 1982 RESULTS (Grade 3)

Scale Categories	Percentage of students obtaining each category for the items		Percentage of students obtaining each category: Average percentage for the items
Section A Objective 1: Writing complete sentences Students were required to rearrange the words to make complete sentences, with correct capitalization and punctuation.			
	1. barked dog Tommy at the	2. opened lady the window has the	
a - Complete sentence	96.2%	94.8%	95.5%
b - Not a complete sentence	2.3	3.1	2.7
Omitted	1.6	2.1	1.9
Section A Objective 2: Capitalization and punctuation of sentences			
w - Capital at the beginning and appropriate end punctuation	86.0	77.4	81.7
x - Sentence does not begin with capital	1.9	2.5	2.2
y - Sentence does not end with appropriate punctuation	8.6	14.8	11.7
z - Sentence does not begin with capital or end with appropriate punctuation	2.0	3.2	2.6
Omitted	1.5	2.1	1.9

SENTENCE WRITING: 1982 RESULTS (Grade 3)

Scale Categories	Percentage of students obtaining each category for the items		Percentage of students obtaining each category: Average percentage for the items
Section B-Objective 1: Writing questions Students were required to change the telling sentences into asking sentences.			
	1. The playground is still wet.	2. The children jumped rope at recess.	
a - Correctly ordered as a question	93.7%	81.9%	87.8%
b - Incorrectly ordered as a question	4.5	15.2	9.9
Omitted	1.9	2.9	2.4
Section B-Objective 2: Capitalization and punctuation of questions			
w - Correctly capitalized and punctuated	79.8	68.6	74.2
x - Capitalization error at the beginning of the sentence	3.3	4.0	3.7
y - Error in end punctuation	12.0	20.3	16.2
z - Error in both beginning capitalization and end punctuation	3.2	4.2	3.7
Omitted	1.8	2.9	2.4

Section C-Objective 1: Subject coordination
Students were required to combine several short sentences to form a sentence with a compound subject, using "and"

Scale Categories	Percentage of students obtaining each category for the items: 1. The books were lying on the table. The pencils were lying on the table.	2. My grandfather came to visit. My uncle came to visit.	Percentage of students obtaining each category: Average percentage for the items
a - Combined using subject coordination	62.9%	66.5%	64.7%
b - Combined using clausal coordination	25.6	24.0	24.8
c - Combined in some other fashion	3.8	3.2	3.5
d - Incorrectly combined	5.8	4.2	5.0
Omitted	2.0	2.1	2.1

Section C-Objective 2: Using commas in a series
Students were required to combine several short sentences to form a sentence using a series.

Scale Categories	3. Today we will study science. Today we will study reading. Today we will study art.	4. The men are going with us. The women are going with us. The children are going with us.	Average percentage for the items
a - Correctly combined using a series with correct punctuation	42.8	35.0	38.9
b - Correctly combined using a series with incorrect internal punctuation (no commas)	7.3	5.9	6.6
c - Correctly combined using and between nominal elements	6.5	8.3	7.4
d - Correctly combined using and between clausal elements	23.1	21.1	22.1
e - Correctly combined in some other fashion	9.0	11.1	10.1
f - Incorrectly combined	8.1	14.5	11.3
Omitted	3.3	4.1	3.7

SENTENCE WRITING: 1982 RESULTS

(Grade 3)

Scale Categories	Percentage of students obtaining each category for the items		Percentage of students obtaining each category: Average percentage for the items
Section D-Objective 1: Using relativization Students were required to combine two short sentences using one of the relative pronouns provided.			
	1. Cary helped the old woman. The old woman was carrying groceries. (who, which, that)	2. We looked across the room. The room filled with students. (who, which, that)	
a - Correctly combined using relativization	67.4	64.9	66.2
b - Correctly combined in some other fashion	3.6	2.4	3.0
c - Incorrectly combined	26.1	29.1	27.6
Omitted	2.9	3.7	3.3
Section D-Objective 2: Using adverbial subordination Students were required to combine two short sentences using one of the conjunctions provided.			
	3. The snake bit Tony. He was teasing it. (and, because, if)	4. Our teacher said we could leave. We had finished our papers. (after, and, without)	
a - Correctly combined using subordination (that is, the two sentences combined using because, if, after, before)	86.1	84.1	85.1
b - Correctly combined using coordination (that is, the two sentences combined using and)	1.4	0.6	1.0
c - Correctly combined in some other fashion	0.6	1.1	0.9
d - Sentences not combined	1.2	0.7	1.0
e - Incorrectly combined	6.6	8.3	7.5
Omitted	4.2	5.2	4.7

CHAPTER 3

GRADE 6 TEST RESULTS

The grade 6 writing test consisted of four parts:

Part I - Spelling: list of 24 spelling words dictated by the teacher.

Part II - Story Writing: written product (narration) in response to a two-sentence story beginning.

Part III - Sentence Writing: sentence production based on combining sets of short sentences.

Part IV - Writing to Explain: written product (process writing) explaining how to do something.

Parts II and IV resulted in a written composition. Both written products were scored according to the same scales and descriptions.

Results of the 1982 assessment are presented in the order specified below. A brief description of the scales and scoring procedures precedes the presentation of results.

WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 42
  General Impression Marking Scale. 42
  Analytic Scales 42
  Story Writing: GIM Results 44
  Story Writing: Analytic Scale Results. 46
  Writing to Explain: GIM Results. 50
  Writing to Explain: Analytic Scale Results . . . 52

WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 56
  Attribute Marking Scales. 56
  Descriptive Scales. 56
  Error Counts. 57
  Story Writing: Attribute and Descriptive Scales/Error Counts. 58
  Writing to Explain: Attribute and Descriptive Scales/Error Counts 62

SPELLING. 66
  Reporting of the 1982 Results. 66

SENTENCE WRITING. 68
  Objectives at the Grade 6 Level. 68
  An Example 69
  Results for the Sentence Writing 69
    Sections A and B. 71
    Section C 72
    Section D 74

WRITTEN PRODUCTS:
GENERAL IMPRESSION MARKING AND ANALYTIC SCALES

On the 1982 Manitoba Writing Assessment, students' written products were evaluated in a number of different ways in order to come to conclusions about the writing performance demonstrated by students. All student writing was evaluated on two types of scales: A General Impression Marking Scale (one score) and Analytic Marking Scales (nine scores). Thus, each student paper was assigned ten different scores for each written product (two each in grades 3, 6, and 9; one in grade 12).

The scales are described in detail in the Manual and Scoring Key at each grade level. A summary is given below.

General Impression Marking Scale (GIM)

General Impression Marking (GIM) is a judgement about the overall quality of a student's writing performance, given the purpose of the writing, the audience to which the writing was addressed, and the expectation for that performance at a grade level. On the provincial tests, two purposes were identified for students at grade levels 3, 6, and 9. The first purpose was to relate a story (narration); the second purpose was to describe or explain how to do something (process writing). At the grade 12 level, students were asked only to write to describe or explain a process. At all grade levels and on all tasks, the audience identified for the students was their class.

The GIM is a quick impressionistic score. Each written product from each student was read by two different raters quickly for a general reaction. No marks were placed on the papers themselves. Each rater assigned a score of 1 to 6 (1 - low, 6 - high). If the scores of the two raters did not differ more than one point (e.g., one rater assigned the paper a 5, the other assigned it a 4), the scores were added together (5 + 4 = 9). That summed score is the GIM score reported for students on the writing task. If the scores of the two raters did differ more than one point (e.g., one rater assigned the paper a 2, the other a 5), the paper was read by an arbitrator. (Arbitrators were scorers chosen because of their experience in doing this type of marking.) In the arbitrated cases, the GIM score for a piece of writing was calculated as follows:

$$\frac{\text{FIRST SCORE} + \text{SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{GIM SCORE}$$

The lowest score a written product could receive was 2 and the highest was 12.

The GIM score attempts to rate compositions in a general, global manner (See Cooper & Odell, 1977 for a complete discussion). Covered by the GIM are any qualities that impress the rater including the quality of thought evidenced in the paper, the effectiveness of the overall organization, the effectiveness of the overall presentation, and the effectiveness of the style and word choice.

Analytic Scales

Like general impression marking, the Analytic Scales are quick, qualitative judgements about the effectiveness of a piece of writing. Analytic Scales, however, are measures of the effectiveness of particular qualities of a paper. For the provincial writing assessment, nine different qualities were

selected for testing, five rhetorical or discourse categories, four mechanics and convention categories. These scales were developed because they seemed to be the qualities most often evaluated in writing by teachers and because they were objectives most often identified as being taught by teachers.

Like the GIM, a paper was assigned a score between 1 and 6 (1 - low, 6 - high), but for each of the nine qualities.

Rhetorical Categories

1. Organization and Sequence
2. Wording and Style
3. Focus or Theme
4. Ideas
5. Choice of Details and Elaboration of Ideas

Mechanics and Conventions Categories

6. Sentence Structure
7. Usage
8. Punctuation and Capitalization
9. Spelling

A scorer, usually different from the GIM scorers, read each written product for the Analytic Scales. Specific descriptions of the qualities associated with high, middle, and low papers are given in each grade level Manual and Scoring Key, as well as examples of papers. (For a more complete discussion of analytic scales in general, refer to Diederich, 1974, and Cooper & Odell, 1977.)

For the provincial sample, every paper was rated on all nine analytic scales by at least one rater. A subset of papers at each grade level was marked analytically by a second rater in order to calculate the amount of agreement or reliability between independent raters.

These two scoring procedures (GIM and Analytic) yielded ten different scores for each written product. At the grade 3, 6, and 9 levels, students wrote two pieces of extended writing so that each student in those grades received 20 different scores (10 for each writing product). At the grade 12 level, the students were asked to write one piece of extended writing so that each student received 10 different scores. This is summarized in Table 2 (p.44).

The GIM and Analytic Score results at the grade 6 level are based on scoring the total provincial sample of 1430 test papers.

TABLE 2

GIM AND ANALYTIC SCORING OF WRITTEN PRODUCTS

Scales	Narrative Writing 3	6	9	12	Descriptive/Expository Writing 3	6	9	12
GIM (Scores range from 2 to 12)	x	x	x	-	x	x	x	x
Analytic (Scores range from 1 to 6)								
1. Organization and sequence	x	x	x	-	x	x	x	x
2. Wording and style (Word Choice)	x	x	x	-	x	x	x	x
3. Focus or theme	x	x	x	-	x	x	x	x
4. Ideas	x	x	x	-	x	x	x	x
5. Choice of Details & Elaboration of Ideas	x	x	x	-	x	x	x	x
6. Sentence Structure	x	x	x	-	x	x	x	x
7. Usage	x	x	x	-	x	x	x	x
8. Punctuation & Capitalization	x	x	x	-	x	x	x	x
9. Spelling	x	x	x	-	x	x	x	x

Story Writing: GIM Results

Students at the grade 6 level were asked to write a story using the story beginning provided as a stimulus (Part II of the test). Teachers were instructed to read the directions aloud to their students. Students were given approvimately 45 minutes to write their stories.

The results of the GIM scoring for the story are presented on page 45. Scores for the two scorers combined ranged from 2 to 12, with scores 2, 3, and 4 described as "Low"; 5, 6, 7, 8, and 9 described as "Middle"; and 10, 11, and 12 described as "High".

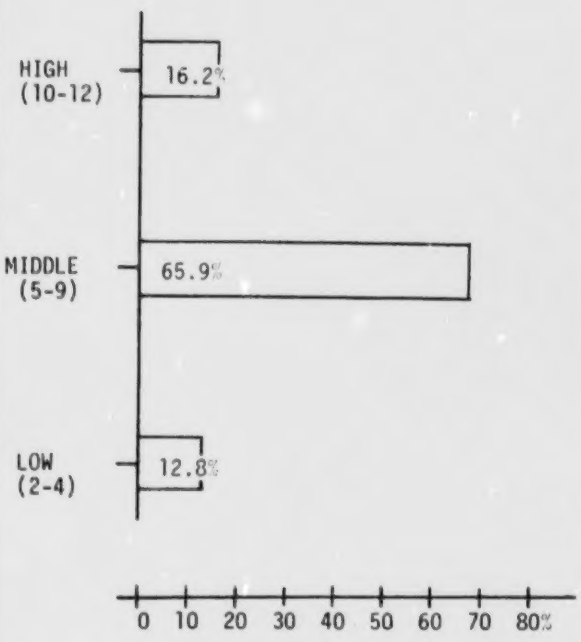
The results are presented in three ways:

1. The percentage of students obtaining each GIM score: e.g., 3.1% of students received a combined GIM score of 12 (both scorers awarded a 6 to these stories).
2. The percentage of students obtaining GIM scores categorized as High, Middle, and Low: e.g., 16.2% of students received combined GIM scores of 10, 11, or 12 (High).
3. The average (mean) performance of students on the GIM scale (and the standard deviation). These are based on the scale of 2 to 12.

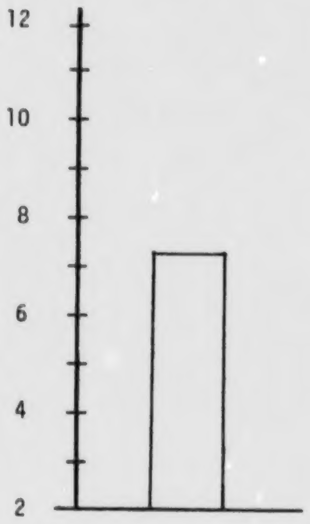
STORY WRITING: 1982 GIM RESULTS (Grade 6)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
	12	3.1%
HIGH	11	5.0
	10	8.1
	9	11.7
	8	11.2
MIDDLE	7	18.0
	6	13.5
	5	11.3
	4	7.7
LOW	3	3.1
	2	2.0
Insufficient Information		4.9

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 7.14

Standard Deviation = 2.33

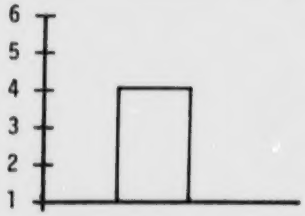
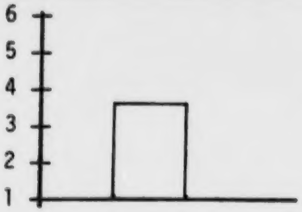
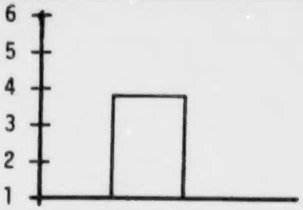
Story Writing: Analytic Scale Results

The results of the analytic scoring for the grade 6 story are presented on pages 47 to 49. All of the scores range on a scale from 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High".

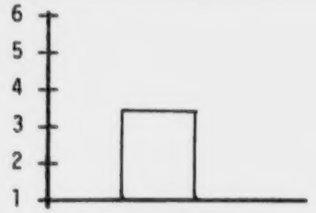
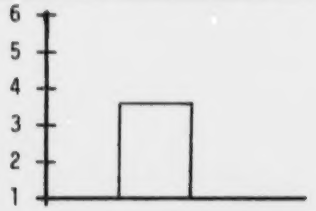
Results are presented in two ways:

1. Percentage of students obtaining each score: e.g., 9.4% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students, and standard deviation, based on the scale of 1 to 6: e.g., the average of students was 4.01 (standard deviation was 1.20) on Organization and Sequence.

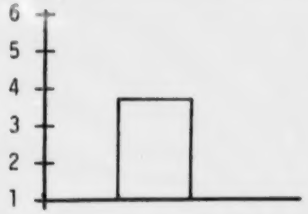
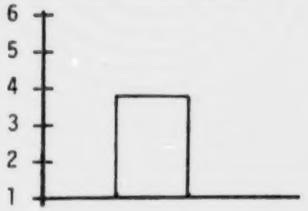
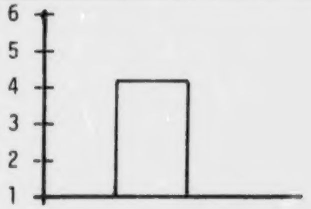
STORY WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	9.4%	
		5	28.3	
	MIDDLE	4	27.4	
		3	22.1	
	LOW	2	9.1	
		1	1.9	
	Insufficient Information		1.8	Average = 4.01 Standard Deviation = 1.20
2. Wording and Style	HIGH	6	6.1	
		5	16.2	
	MIDDLE	4	27.9	
		3	35.0	
	LOW	2	12.0	
		1	1.3	
	Insufficient Information		1.6	Average = 3.65 Standard Deviation = 1.12
3. Focus or Theme	HIGH	6	7.3	
		5	22.2	
	MIDDLE	4	29.2	
		3	23.0	
	LOW	2	11.7	
		1	3.4	
	Insufficient Information		3.1	Average = 3.80 Standard Deviation = 1.24

STORY WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	6.9%	
		5	17.5	
	MIDDLE	4	27.6	
		3	24.5	
	LOW	2	14.3	
		1	2.4	
	Insufficient Information		6.8	Average = 3.69 Standard Deviation = 1.22
5. Choice of Details and Elaboration of Ideas		6	5.0	
	HIGH	5	16.1	
		4	24.3	
	MIDDLE	3	26.7	
	LOW	2	18.7	
		1	3.5	
	Insufficient Information		5.8	Average = 3.49 Standard Deviation = 1.24
6. Sentence Structure	HIGH	6	4.5	
		5	19.9	
	MIDDLE	4	26.6	
		3	26.9	
		2	14.4	
	LOW	1	5.1	
	Insufficient Information		2.6	Average = 3.57 Standard Deviation = 1.24

STORY WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	5.5	 Average = 3.77 Standard Deviation = 1.24
		5	24.8	
	MIDDLE	4	26.9	
		3	21.3	
	LOW	2	13.1	
		1	3.4	
	Insufficient Information		5.1	
8. Punctuation and Capitalization	HIGH	6	8.4	 Average = 3.81 Standard Deviation = 1.37
		5	26.6	
	MIDDLE	4	24.3	
		3	18.7	
	LOW	2	12.6	
		1	6.0	
	Insufficient Information		3.5	
9. Spelling	HIGH	6	13.1	 Average = 4.11 Standard Deviation = 1.35
		5	28.9	
	MIDDLE	4	20.8	
		3	14.9	
	LOW	2	9.9	
		1	3.5	
	Insufficient Information		9.0	

Writing to Explain: GIM Results

Students at the grade 6 level were asked to explain how to do something, using as suggested topics: How to make something, how to shop for and buy something, or how to do a job (Part IV of the test). Teachers were instructed to read the directions aloud to their students. Students were given approximately 50 minutes to write their directions.

These student compositions were scored using the same GIM and analytic scales as were used for the stories (see pages 42-44).

The results of the GIM scoring are presented on page 51. Scores for the two scorers combined ranged from 2 to 12, with scores 2, 3, and 4 categorized as "Low; 5, 6, 7, 8, and 9 as "Middle"; and 10, 11, and 12 as "High".

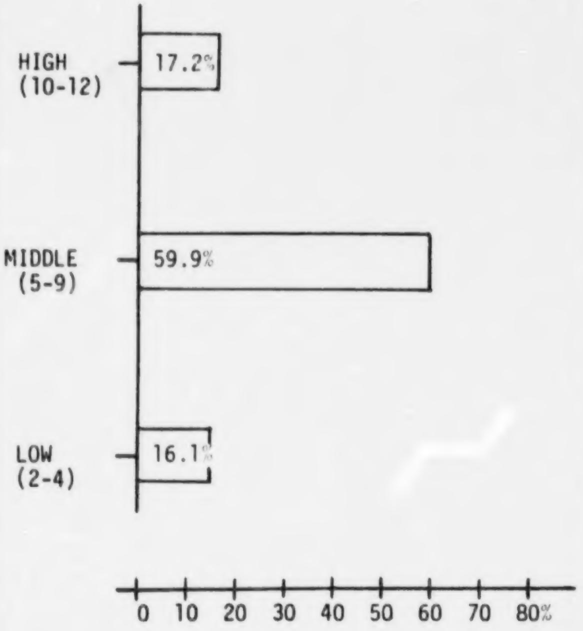
The results are depicted in three ways:

1. The percentage of students obtaining each GIM score: e.g., 2.5% of students received a combined GIM score of 12.
2. The percentage of students obtaining GIM scores categorized as High, Medium, and Low: e.g., 17.2% of students received scores in the High category (GIM score of 10, 11, or 12).
3. The average (mean) performance of students on the GIM scale (and the standard deviation).

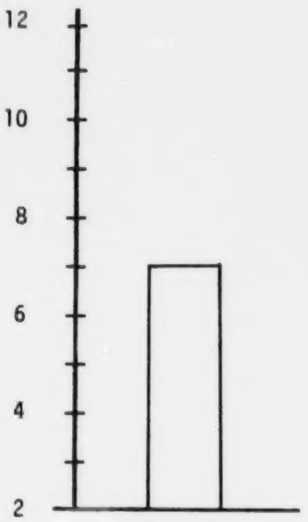
WRITING TO EXPLAIN: 1982 GIM RESULTS (Grade 6)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
	12	2.5%
HIGH	11	6.6
	10	8.1
	9	12.1
	8	9.8
MIDDLE	7	15.7
	6	11.0
	5	11.3
	4	6.7
LOW	3	5.4
	2	4.0
Insufficient Information		6.5

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 7.03

Standard Deviation = 2.54

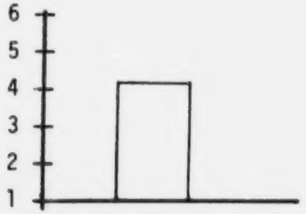
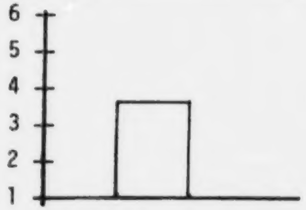
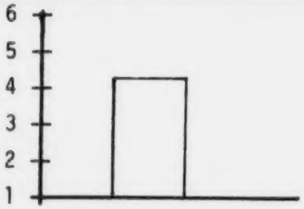
Writing to Explain: Analytic Scale Results

The results of the analytic scoring for the grade 6 process writing are presented on pages 53 to 55. All of the analytic scores are on a scale of 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High". These scales are identical to those used in evaluating the narrative writing (see pages 42 to 44 for a detailed description).

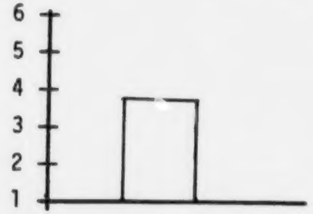
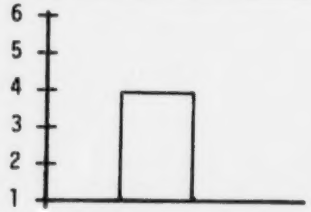
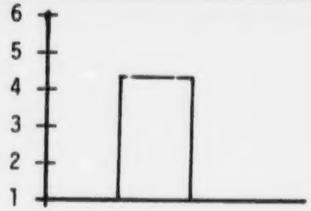
Results for each analytic scale are depicted in two ways:

1. Percentage of students obtaining each score: e.g., 13.6% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students on each analytic scale (and standard deviation) based on the scale from 1 to 6: e.g., the average performance was 4.07 (standard deviation was 1.29) on Organization and Sequence.

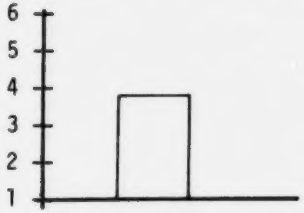
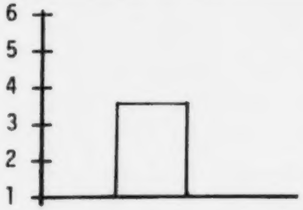
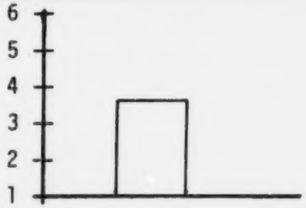
WRITING TO EXPLAIN: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	13.6	
		5	26.6	
	MIDDLE	4	24.8	
		3	20.0	
	LOW	2	9.3	
		1	2.7	
	Insufficient Information		3.1	Average = 4.07 Standard Deviation = 1.29
2. Wording and Style	HIGH	6	5.7	
		5	19.1	
	MIDDLE	4	30.0	
		3	28.0	
	LOW	2	11.5	
		1	2.4	
	Insufficient Information		3.3	Average = 3.71 Standard Deviation = 1.16
3. Focus or Theme	HIGH	6	14.8	
		5	28.9	
	MIDDLE	4	24.1	
		3	15.4	
	LOW	2	8.3	
		1	4.5	
	Insufficient Information		4.0	Average = 4.13 Standard Deviation = 1.36

WRITING TO EXPLAIN: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	9.0%	
		5	24.5	
	MIDDLE	4	25.9	
		3	21.6	
	LOW	2	11.7	
		1	3.5	
	Insufficient Information		3.7	Average = 3.86 Standard Deviation = 1.28
8. Punctuation and Capitalization	HIGH	6	11.0	
		5	30.7	
	MIDDLE	4	22.2	
		3	17.0	
	LOW	2	9.2	
		1	6.8	
	Insufficient Information		3.1	Average = 3.97 Standard Deviation = 1.40
9. Spelling	HIGH	6	18.3	
		5	29.6	
	MIDDLE	4	20.3	
		3	14.8	
	LOW	2	8.9	
		1	4.3	
	Insufficient Information		3.9	Average = 4.22 Standard Deviation = 1.40

WRITING TO EXPLAIN: 1982 ANALYTIC SCALE RESULTS (Grade 6)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	9.4%	
		5	22.1	
	MIDDLE	4	27.6	
		3	22.0	
	LOW	2	11.3	
		1	3.4	
	Insufficient Information		4.3	Average = 3.86 Standard Deviation = 1.28
5. Choice of Details and Elaboration of Ideas	HIGH	6	7.5	
		5	16.4	
	MIDDLE	4	24.3	
		3	25.5	
	LOW	2	17.5	
		1	5.7	
	Insufficient Information		3.0	Average = 3.52 Standard Deviation = 1.33
6. Sentence Structure	HIGH	6	5.3	
		5	19.4	
	MIDDLE	4	27.8	
		3	25.9	
	LOW	2	14.1	
		1	4.5	
	Insufficient Information		2.9	Average = 3.61 Standard Deviation = 1.24

WRITTEN PRODUCTS:

ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS

In addition to the qualitative judgements made on each paper using the GIM and Analytic Scales, a randomly selected sub-sample of papers was examined for more intensive structural, linguistic, attribute, and error descriptions. These in-depth analyses were carried out by the same group of raters employed for the other marking tasks. A total of 63 student papers were selected at the grade 6 level for this intensive scoring.

The rationale for these types of analyses was to make assessments regarding the kinds of structures students used in their actual writing and the kinds of errors most frequently made by students in actual writing situations. Also, they were conducted in order to draw some conclusions about the number of students demonstrating serious problems in the structural aspects of their writing. The choice of structures, errors, and attributes was based on the survey of writing done by Hunt (See Cooper & Odell, 1977), on the work of Evans (1979) in the Ontario assessment, and on the objectives identified by practicing teachers as the ones of greatest interest and use to them in their classrooms.

Attribute Marking Scales

Seven different attribute marking scales were identified as being of interest to teachers. The scales themselves were evaluations about whether or not a student was experiencing serious (in need of remedial help) difficulty in particular areas. For each paper evaluated, the rater was asked to rate the composition on the seven attributes, each on a three-point scale:

3 - No errors;
2 - An occasional error, but not rated as a serious problem;
1 - Serious problem (Student is in need of special remedial help);
X - Insufficient information (Structures not used or insufficient writing for an evaluation to be made).

Five of the seven areas were evaluated at all grade levels. The last two (use of contractions and use of quotation marks in direct discourse) were evaluated only at grade levels 3 and 6. The seven attributes were:

1. Run-on sentences
2. Use of complete sentences
3. Paragraphing
4. Internal punctuation
5. Capitalization
6. Use of contractions
7. Use of quotation marks in direct discourse

Descriptive Scales

The descriptive scales were counts of particular structures used by students. The following counts were made for each student composition in grades 3, 6, and 9. Grade 12 compositions were tallied on all but the last one.

1. Total number of words written
2. Total number of T-units (independent clauses) written
3. Total number of dependent clauses written
4. Total number of adverbial clauses written

5. Total number of relative clauses written
6. Total number of noun clauses written
7. Total number of modifiers (not assessed due to the large number)
8. Total number of instances of direct discourse

The particular counts were chosen because they seem closely related both to composition quality and to syntactic fluency (See Hunt in Cooper & Odell, 1977). The purpose was to judge whether or not students easily and fluently used these structures in their actual writing.

From these counts, a number of transformed scores were calculated to produce twelve indices of writing maturity in students. In general, the more fluent, mature writers will demonstrate higher scores on each of these indices, whereas weaker, less mature writers will score lower. These scores should be viewed, however, only as indicators of how students handle the syntax and grammar of the language in free writing situations. The question answered by these data is: Do students use these particular structures in their own writing? The question of whether or not students can use these structures is left to be answered by the sentence writing tasks (Part III).

The scores calculated included:

Words per T-unit
Words per clause
Clauses per T-unit
Dependent clauses per T-unit
Adverbial clauses per T-unit
Adverbial clauses per 100 words
Relative clauses per T-unit
Relative clauses per 100 words
Noun clauses per T-unit
Noun clauses per 100 words
Direct discourse per T-unit
Direct discourse per 100 words

Error Counts

Error counts, though tedious to do, can give information about how well students avoid error in actual writing situations. These counts are realistic alternatives to the use of objective (particularly multiple-choice type) items in evaluating writing. Error counts can give greater information about the choices that students are making in actual writing tasks.

Each of the error categories are calculated in these ways:

average number of errors per composition
average number of errors per T-unit (i.e. per independent clause), and
average number of errors per 100 running words

The error counts obtained were:

1. Number of sentence fragments
2. Number of incorrectly placed modifiers
3. Number of sentences not beginning with a capital letter
4. Number of proper nouns not capitalized
5. Number of sentences not ending with appropriate punctuation
6. Number of internal comma errors
7. Number of faulty subject-verb agreement
8. Number of verb tense errors
9. Number of word order errors

Similarly to the descriptive scales a number of scores were calculated from the error counts. But, opposite to the descriptive scales, higher scores on the scale reflect greater error rates.

The calculated scores were:

Sentence fragments per T-unit
Sentence fragments per 100 words
Incorrectly placed modifiers per T-unit
Incorrectly placed modifiers per 100 words
Beginning capitalization errors per T-unit
Beginning capitalization errors per 100 words
Proper nouns not capitalized per T-unit
Proper nouns not capitalized per 100 words
End punctuation errors per T-unit
End punctuation errors per 100 words
Comma errors per T-unit
Comma errors per 100 words
Faulty agreement per T-unit
Faulty agreement per 100 words
Verb tense errors per T-unit
Verb tense errors per 100 words
Word order errors per T-unit
Word order errors per 100 words

Story Writing: Attribute and Descriptive Scales/Error Counts

The results for the grade 6 stories (Part II) are presented for:

The 7 attribute scales (p. 59),

The 8 descriptive scales (p. 60),

The 12 indices calculated from the descriptive scales (p. 60),

The 9 error counts (p. 61), and

The 18 indices calculated from the error counts (p. 61).

Since there are a large number of scales, the results have been tabulated in a compacted form in the following tables.

The average (mean) and standard deviation are reported for each scale and index. For the attribute scales, the percentage of students obtaining each score is reported: e.g., 42.9% of students obtained a score of 3 (no errors) on the attribute scale "Run-on sentences", and the average is 2.38 on the scale of 1 to 3 (standard deviation is 0.61).

STORY WRITING: 1982 ATTRIBUTE SCALES RESULTS (Grade 6)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
1. Run-on sentences				
No errors	3	42.9%		
An occasional error	2	47.6	2.38	0.61
Serious problem	1	6.3		
Insufficient information		3.2		
2. Use of complete sentences				
No errors	3	44.4		
An occasional error	2	49.2	2.45	0.53
Serious problem	1	1.6		
Insufficient information		4.8		
3. Paragraphing				
No errors	3	38.1		
An occasional error	2	28.6	2.14	0.83
Serious problem	1	25.4		
Insufficient information		8.0		
4. Internal punctuation				
No errors	3	22.2		
An occasional error	2	68.3	2.16	0.52
Serious problem	1	6.3		
Insufficient information		3.2		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
5. Capitalization				
No errors	3	52.4%		
An occasional error	2	34.9	2.44	0.67
Serious problem	1	9.5		
Insufficient information		3.2		
6. Use of contractions*				
No errors	3	38.1		
An occasional error	2	15.9	2.61	0.60
Serious problem	1	3.2		
Insufficient information		42.9		
7. Use of quotation marks in direct discourse*				
No errors	3	27.0		
An occasional error	2	19.0	2.79	0.78
Serious problem	1	11.0		
Insufficient information		42.9		

*Many students did not use these particular features in their story writing; this was categorized as "Insufficient information".

STORY WRITING: 1982 DESCRIPTIVE SCALE RESULTS (Grade 6)

	Descriptive Scale (Actual counts)	Average Number	
		Mean	Standard Deviation
1.	Words	265.13	93.38
2.	T-units	27.76	12.73
3.	Dependent Clauses	9.23	5.61
4.	Adverbial clauses	5.16	3.29
5.	Relative clauses	0.82	1.41
6.	Noun clauses	3.07	3.66
7.	Modifiers	(not counted)	
8.	Instances of direct discourse	3.45	5.90

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	10.03	3.09
Words per clause*	7.31	1.62
Clauses per T-unit	1.37	0.26
Dependent clauses per T-unit	0.37	0.26
Adverbial clauses per T-unit	0.21	0.18
Adverbial clauses per 100 words	2.00	1.15
Relative clauses per T-unit	0.033	0.052
Relative clauses per 100 words	0.30	0.49
Noun clauses per T-unit	0.11	0.13
Noun clauses per 100 words	1.08	1.21
Direct discourse per T-unit	0.10	0.14
Direct discourse per 100 words	1.15	1.70

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

NARRATIVE WRITING: ERROR COUNTS (Grade 6)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.020	0.062
Sentence fragments per 100 words	0.22	0.73
Incorrectly placed modifiers per T-unit	0.010	0.022
Incorrectly placed modifiers per 100 words	0.11	0.24
Beginning capitalization errors per T-unit	0.12	0.18
Beginning capitalization errors per 100 words	1.29	2.06
Proper nouns not capitalized per T-unit	.008	0.035
Proper nouns not capitalized per 100 words	0.11	0.55
End punctuation errors per T-unit	0.11	0.15
End punctuation errors per 100 words	1.25	1.79
Comma errors per T-unit	0.12	0.12
Comma errors per 100 words	1.28	1.44
Faulty agreement per T-unit	0.011	0.032
Faulty agreement per 100 words	0.11	0.27
Verb tense errors per T-unit	0.016	0.034
Verb tense errors per 100 words	0.19	0.44
Word order errors per T-unit	0.012	0.028
Word order errors per 100 words	0.13	0.32

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Standard Deviation
1. Sentence fragments	0	69.8%	0.68	2.59
	1	22.2		
	2 or more	8.0		
2. Incorrectly placed modifiers	0	79.4	0.32	0.74
	1	9.5		
	2 or more	11.1		
3. Sentences not beginning with Capital letters	0	50.8	3.68	6.58
	1	9.5		
	2 or more	39.7		
4. Proper nouns not capitalized	0	90.5	0.40	2.43
	1	3.2		
	2 or more	6.3		
5. Sentences not ending with appropriate punctuation	0	41.3	3.63	6.10
	1	14.3		
	2 or more	44.4		
6. Internal comma errors	0	25.4	3.61	4.98
	1	11.1		
	2 or more	63.5		
7. Instances of faulty subject/verb agreement	0	81.0	0.25	0.75
	1	14.3		
	2 or more	4.7		
8. Verb tense errors	0	71.4	0.57	1.73
	1	15.9		
	2 or more	12.7		
9. Word order errors	0	77.8	0.35	0.94
	1	14.3		
	2 or more	7.9		

Writing to Explain: Attribute and Descriptive Scales/Error Counts

The results for the process writing part of the grade 6 test (Part IV) are presented for the scales and calculated indices. The scales are identical to those used for the stories (for a description see pages 56 to 58):

The 7 attribute scales (p. 63),
The 8 descriptive scales (p. 64),
The 12 indices calculated from the descriptive scales (p. 64),
The 9 error counts (p. 65), and
The 18 indices calculated from the error counts (p. 65).

Due to the large number of scales and indices the results have been tabulated in a compacted form in the following tables. The average (mean) and standard deviation are reported for each scale and index. As well, for the attribute scales, the percentage of students obtaining each score is reported: e.g., 36.5% of students obtained a score of 3 (no errors) on the attribute scale "Run-on Sentences", and the average is 2.38 on the scale of 1 to 3 (standard deviation 0.64).

The results are based on the sub sample of 63 students. Ten students of the 63 did not produce any writing for this part of the test (Part IV); the papers of these students were excluded from the calculation of the average and are categorized as "Insufficient information".

WRITING TO EXPLAIN: 1982 ATTRIBUTE SCALES RESULTS (Grade 6)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
1. Run-on sentences				
No errors	3	36.5		
An occasional error	2	36.5	2.38	0.64
Serious problem	1	6.3		
Insufficient information		20.6		
2. Use of complete sentences				
No errors	3	34.9		
An occasional error	2	41.3	2.40	0.57
Serious problem	1	3.2		
Insufficient information		20.6		
3. Paragraphing				
No errors	3	44.4		
An occasional error	2	17.5	2.52	0.70
Serious problem	1	7.9		
Insufficient information		20.6		
4. Internal punctuation				
No errors	3	22.2		
An occasional error	2	49.2	2.23	0.56
Serious problem	1	4.8		
Insufficient information		23.8		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
5. Capitalization				
No errors	3	34.9		
An occasional error	2	36.5	2.37	0.64
Serious problem	1	6.3		
Insufficient information		22.2		
6. Use of contractions*				
No errors	3	12.7		
An occasional error	2	9.5	2.29	0.77
Serious problem	1	4.8		
Insufficient information		73.0		
7. Use of quotation marks in direct discourse*				
No errors	3	9.5		
An occasional error	2	1.6	-	-
Serious problem	1	0.0		
Insufficient information		88.9		

*Many students did not use these particular features in their writing; this was categorized as "Insufficient information".

WRITING TO EXPLAIN: 1982 DESCRIPTIVE SCALE RESULTS

(Grade 6)

Descriptive Scale (Actual counts)	Average Number Mean	Average Number Standard Deviation
1. Words	165.77	86.19
2. T-units	14.49	8.50
3. Dependent Clauses	4.55	3.39
4. Adverbial clauses	3.59	2.86
5. Relative clauses	0.45	0.77
6. Noun clauses	0.49	1.10
7. Modifiers	(not counted)	
8. Instances of direct discourse	0.09	0.35

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	11.99	2.83
Words per clause*	8.96	1.89
Clauses per T-unit	1.35	0.24
Dependent clauses per T-unit	0.35	0.24
Adverbial clauses per T-unit	0.28	0.22
Adverbial clauses per 100 words	2.24	1.62
Relative clauses per T-unit	0.034	0.062
Relative clauses per 100 words	0.26	0.44
Noun clauses per T-unit	0.036	0.085
Noun clauses per 100 words	0.28	0.62
Direct discourse per T-unit	0.006	0.024
Direct discourse per 100 words	0.058	0.269

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

WRITING TO EXPLAIN: 1982 ERROR COUNTS (Grade 6)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.031	0.069
Sentence fragments per 100 words	0.29	0.61
Incorrectly placed modifiers per T-unit	0.008	0.023
Incorrectly placed modifiers per 100 words	0.071	0.216
Beginning capitalization errors per T-unit	0.15	0.23
Beginning capitalization errors per 100 words	1.40	2.49
Proper nouns not capitalized per T-unit	0.006	0.029
Proper nouns not capitalized per 100 words	0.046	0.233
End punctuation errors per T-unit	0.14	0.19
End punctuation errors per 100 words	1.37	2.32
Comma errors per T-unit	0.11	0.13
Comma errors per 100 words	0.99	1.18
Faulty agreement per T-unit	0.001	0.007
Faulty agreement per 100 words	0.009	0.068
Verb tense errors per T-unit	0.003	0.010
Verb tense errors per 100 words	0.023	0.095
Word order errors per T-unit	0.017	0.049
Word order errors per 100 words	0.14	0.42

Error Counts		Percentage of Students Obtaining each Count	Average Error Count	
			Mean	Standard Deviation
1. Sentence fragments	0	71.7%	0.47	1.12
	1	20.8		
	2 or more	7.6		
2. Incorrectly placed modifiers	0	88.7	0.13	0.39
	1	9.4		
	2 or more	1.9		
3. Sentences not beginning with Capital letters	0	41.5	2.28	4.25
	1	18.9		
	2 or more	39.6		
4. Proper nouns not capitalized	0	96.2	0.06	0.31
	1	1.9		
	2 or more	1.9		
5. Sentences not ending with appropriate punctuation	0	37.7	2.06	3.50
	1	24.5		
	2 or more	37.8		
6. Internal comma errors	0	47.2	1.59	2.12
	1	11.3		
	2 or more	41.5		
7. Instances of faulty subject/verb agreement	0	98.1	0.02	0.14
	1	1.9		
	2 or more	0.0		
8. Verb tense errors	0	94.3	0.06	0.23
	1	5.7		
	2 or more	0.0		
9. Word order errors	0	86.8	0.23	0.70
	1	7.5		
	2 or more	5.7		

SPELLING

The purpose of the spelling assessment on the 1982 Writing Assessment Program was to determine if students could spell correctly the words they actually used in their writing as well as the words they were likely to need in their writing. Spelling in the actual written products was assessed by one of the analytic scales (Analytic Scales, p. 42). In order to assess the performance of students on the spelling of words they were likely to need in their general writing (but may not have used in the writing tasks provided on the test), words were randomly selected from Ves Thomas' Canadian Word Lists (1979). This particular set of word lists assigns words to different grade levels based on the frequency of the appearance of these words in the writing of students at a particular grade level. For example, the grade 6 list is composed of words that a high number of Canadian grade 6 students use in their own writing.

In addition to the words selected from the Thomas Canadian Word Lists, a number of words were selected from the 1979 Manitoba Writing Assessment Program so that comparisons between student performance in 1979 and 1982 could be made. These 1979 words were originally selected not because they should be in students' writing vocabularies but because they conformed to particular patterns and appeared in grade level reading materials: that is, they were words that students could be expected to recognize and pronounce (but not necessarily be able to use or spell in writing). The rationale for the selection of the 1982 spelling words seemed stronger in that it reflected words that students were likely to use in composing rather than in reading.

The spelling test used in the 1982 assessment contained two separate tests (Part I of the test). One set of items (those marked with a * in the Manual and Scoring Key) was taken from the 1979 Manitoba Writing Assessment. The total number of these words is 8.

The balance of the words were randomly selected from the Thomas list at the grade 6 level. There were 16 words obtained in this way for the 1982 component.

Two scores were given on the test, one for the items from 1979 (Total possible = 8), another for the items selected for 1982 (Total possible = 16). Total overall performance (Total possible = 24) was determined by adding the two scores together. The results are based on the provincial sample of 1430 grade 6 students.

Reporting of the 1982 Results

The average (mean) score is reported for the scores based on the 8 items, on the 16 items, and on the 24 items (combination of 8 and 16 items). A frequency distribution indicating the percentage of students obtaining each score is presented (e.g., 23.1% of students spelled 21 or 22 of the words correctly). Finally, the percentage of students spelling each word correctly is indicated on page 67 (e.g., 85.1% of students spelled "crowd" correctly).

SUMMARY OF THE 1982 SPELLING RESULTS (Grade 6)

Score Correct out of Total Number of Words (25)	Percentage of Students Obtaining Each Score
23-24	28.8%
21-22	23.1
19-20	15.6
17-18	9.2
15-16	7.4
13-14	4.8
11-12	3.4
9-10	2.3
7-8	2.0
5-6	1.7
3-4	0.4
1-2	0.2
0	1.0

Average Score Correct	19.11
Standard Deviation	4.96
Average Percentage Correct	79.6%

Words Selected From the 1979 Assessment		Words Selected For the 1982 Assessment	
Items	Percentage of 1982 students obtaining item correct	Items	Percentage of 1982 students obtaining item correct
7. questions	84.7%	1. crowd	85.1%
8. information	85.7	2. shouted	93.6
9. accident	65.9	3. fighting	95.2
11. beautiful	73.5	4. behind	93.8
15. straight	76.5	5. except	79.0
18. grabbed	70.1	6. Thursday	77.7
21. finally	72.6	10. helicopter	63.0
23. piece	72.8	12. yourself	95.5
		13. station	90.6
		14. thought	93.3
		16. mammal	55.3
		17. wonderful	88.7
		19. tied	88.1
		20. gigantic	70.1
		22. wasn't	84.3
		24. invisible	55.9

Average Correct	6.02	Average Correct	13.09
Standard Deviation	2.17	Standard Deviation	3.07
Average Percentage Correct	75.2%	Average Percentage Correct	81.8%

SENTENCE WRITING

A number of sentence writing tasks were designed to assess performance on selected structural, syntactic, usage, and mechanics objectives. The format of the sentence writing tasks was chosen for five reasons:

> It closely resembled the kind of thinking and shaping a writer probably goes through when composing.
>
> It allowed for a variety of answers which could be labeled "correct" so that a student was not limited to a single correct answer and could create the most effective sentence(s) for the situation and information provided.
>
> The structurally correct answers in most cases could be generally categorized from least mature (or fluent) to most mature (or fluent).
>
> The open format centred on what students could do in particular circumstances rather than on what they could not do.
>
> The format could allow for the assessment of whether students could use particular structures.

Performance on the sentence writing tasks could be compared with the descriptive scales in the actual writing tasks in order to determine if students in fact did use the targeted structures in free writing.

Although a number of objectives (or targeted structures) were common across all four grade levels, the particular structures or conventions assessed at each grade level were based on objectives that practicing teachers felt were applicable at that level.

Objectives at the Grade 6 level

Objectives at the grade 6 level are listed below; the sections comprised Part III of the student test.

Section A - Objective: Use of participles
Section B - Objective: Use of relativization
Section C - Objective 1: Use of predicate coordination
Section C - Objective 2: Capitalization and punctuation in a sentence
Section C - Objective 3: Use of commas in a series
Section D - Objective 1: Use of adverbial subordination
Section D - Objective 2: Capitalization and punctuation in a sentence

Several sentences were required for each section but some sentences were scored for two objectives.

The objectives fall into two basic areas: Capitalization and punctuation conventions (C2, C3, D2), and syntactic maturity (A, B, C1, D1). All students should have some concept about both of these areas.

Student sentences were scored according to several categories, denoted a, b, c, etc. The most careful, mature writers would fall into category "a" of each objective. Less mature writers might have a profile of mixed performance: e.g., demonstrate sentence sense and knowledge of punctuation, but not demonstrate the ability to capitalize correctly or use the relatively mature structures such as adverbial subordination. The fact that a student did not consistently answer with the most mature structure does not necessarily mean that she/he does not know how to employ the structure or convention, rather that the structure or convention is not readily accessible in the written mode. Weaker students might demonstrate little fluency in the structures they write or choose. The tasks presented in the assessment allow for divergent answers so that the teacher can assess the level of maturity or fluency of the student.

An Example

For Item 1 of Section C, Objective 1, the student was asked to combine the following two sentences using the word "and":

> The teacher handed back our spelling papers.
> The teacher told us we had done excellent work.

The most fluent writer probably responded with an answer like:

> The teacher handed back our spelling papers and told us we had done excellent work.

This answer would be categorized a: Combined used predicate coordination.

The somewhat less fluent writer might have responded with an answer such as:

> The teacher handed back our spelling papers and she told us we had done excellent work.

This answer would be categorized b: Correctly combined by joining the two sentences with the word "and". This answer is still structurally correct and grammatical; it is simply not as efficient or fluent as the first answer. The categorization scheme for this objective allows the teacher to judge a structure as "correct" but "less mature".

A less fluent writer might have chosen to put the sentences together, but keep them as separate thoughts and sentences, such as:

> The teacher handed back our spelling papers. And the teacher told us we had done excellent work.

This answer would be categorized c: Two complete sentences used. Again, this is a structurally correct and grammatical answer, but it is certainly not as mature as answers categorized a or b.

Another writer might have answered:

> When the teacher handed back our spelling papers, she told us we had done excellent work.

The writer has clearly not followed directions; however, the purpose of the writing assessment is not to evaluate direction-following behavior, but writing behavior. The student has written a structurally correct, grammatical, and meaningful sentence, and should be given credit for it. This answer, in fact, might be thought of as more mature than any previously cited. This answer would be categorized as d: Correctly combined in some other fashion. The maturity or fluency level here is not apparent from the category description, but the scheme does give the writer credit of meaningful writing.

The final category, e: Incorrectly combined, was only used if the student wrote an answer that did not make sense or was not grammatical.

The complete categories and example of answers are given in the grade 6 Manual and Scoring Key.

Results for the Sentence Writing

At the grade 6 level students were asked to do two or more sentence writing tasks (items) for each of the objectives. Some of the sentences in sections C and D were also scored for a second objective regarding capitalization and punctuation.

The results are presented for each sentence and for each objective. Since the scoring categories for each objective were not based on a numerical scale averages cannot be used. Instead reported are the percentages of students whose sentences were categorized in each of the scoring categories (the first scoring category is always the targeted structure of the objective). Also reported is the average of the percentages for a particular category for the two items pertaining to an objective. For example, in Section A 57.6% of student sentences were categorized as "a: Correctly combined using participles with correctly placed modifiers" for Item 1 and 40.6% for Item 2. On average, 49.1% of student sentences were categorized as "a" for the two sentences combined and thus for the objective: Use of participles.

Results are based on the total student sample of 1430 test papers.

SENTENCE WRITING: 1982 RESULTS (Grade 6)

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section A-Objective: Use of participles			
	1. The cattle stood under the oak tree. The cattle were shivering.	2. The home team won the game. The home team was proud. The game was exciting.	
a - Correctly combined using participles with correctly placed modifiers	57.6%	40.6%	49.1%
b - Correctly combined in some other fashion	22.0	21.6	21.8
c - Incorrectly placed modifiers	12.1	8.0	10.1
d - Incorrectly combined	5.6	[illegible]	14.9
Omitted	2.8	[illegible]	4.2

Scale Categories	Percentage of students obtaining each category for the Items			Percentage of students obtaining each category: Average percentage for the items
Section B-Objective: Use of relativization				
	1. The old man ate an apple. The old man sat on a bench.	2. The young girl sang a song. The girl's hair was blonde.	3. The alligator began swimming toward the canoe. The alligator had looked like a log.	
a - Correctly combined using relativization	46.1%	37.1%	46.7%	43.3%
b - Correctly combined in some other fashion	8.7	8.7	7.8	8.4
c - Incorrectly combined	42.2	50.2	41.3	44.6
Omitted	3.0	4.0	4.1	3.7

SENTENCE WRITING: 1982 RESULTS

(Grade 6)

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section C-Objective 1: Use of predicate coordination			
	1. The teacher handed back our spelling papers. The teacher told us we had done excellent work.	2. The workmen plowed the highway. The workmen sprinkled sand on the highway.	
a - Correctly combined using predicate coordination	81.7%	75.0%	78.4%
b - Correctly combined by joining the two sentences with the word "and"	10.2%	17.3	13.8
c - Two complete sentences used	0.7	0.7	0.7
d - Correctly combined in some other fashion	1.9	1.4	1.7
e - Incorrectly combined	2.7	2.0	2.4
Omitted	2.8	3.5	3.2
Section C-Objective 2: Capitalization and Punctuation			
w - No capitalization or punctuation error	83.6%	80.1%	81.9%
x - Student did not capitalize the beginning of the sentence	0.9	1.0	1.0
y - Student did not use end punctuation	10.4	14.1	12.3
z - Student did not use either capital letter at the beginning of the sentence or end punctuation	0.7	0.8	0.8
Omitted	4.4	4.1	4.3

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section C-Objective 3: Use of commas in a series			
	3. Charles wanted to go on the field trip. Anna wanted to go on the field trip. Connie wanted to go on the field trip.	4. I looked for you at school. I looked for you at the playground. I looked for you at home.	
a - Combined in a series with no internal punctuation error	62.5%	55.0%	58.8%
b - Combined in a series with internal punctuation error (no commas)	4.5	8.2	6.4
c - Combined using "and" between elements.	5.1	10.1	7.6
d - Combined by joining the three sentences with the word "and"	15.5	13.0	14.3
e - Correctly combined in some other fashion	4.8	5.5	5.2
f - Incorrectly combined	3.2	3.4	3.3
Omitted	4.4	4.7	4.6

SENTENCE WRITING: 1982 RESULTS (Grade 6)

Scale Categories	Percentage of students obtaining each category for the Items			Percentage of students obtaining each category: Average percentage for the items
Section D-Objective 1: Adverbial subordination				
	1. The baby was crying. He had broken his toy.	2. Charles and Alice caught the bus home from school. They helped their teacher clean up the classroom.	3. Carl fell off his bicycle. He was showing off.	
a - Subordination used	88.4%	78.7%	86.0%	84.4%
b - Coordination used	1.8	5.9	2.8	3.5
c - Correctly combined in some other fashion	0.7	1.3	0.8	0.9
d - Incorrectly combined	3.9	5.7	2.4	4.0
Omitted	5.2	8.3	7.9	7.1
Section D-Objective 2: Capitalization and Punctuation				
v - No capitalization or punctuation error	76.3%	69.4%	75.1%	73.6%
w - Student did not capitalize beginning of the sentence	0.7	0.8	0.4	0.6
x - Student did not use end punctuation	14.0	15.9	10.1	13.3
y - Student did not use either capital letter at the beginning of the sentence or end punctuation	0.5	0.4	0.3	0.4
z - Other capitalization/punctuation error	3.6	4.8	4.2	4.2
Omitted	5.0	8.6	9.8	7.8

CHAPTER 4

GRADE 9 TEST RESULTS

The grade 9 writing test consisted of four parts:

Part I - Spelling: list of 26 spelling words dictated by the teacher.

Part II - Narrative Writing: written product (narration) in response to a three-sentence story beginning.

Part III - Sentence Writing: sentence production based on combining sets of short sentences.

Part IV - Process Writing: written product (process writing) explaining how to do something.

Parts II and IV resulted in a written composition. Both written products were scored according to the same scales and descriptions.

Results of the 1982 assessment are presented in the order specified below. A brief description of the scales and scoring procedures precedes the presentation of results.

WRITTEN PRODUCTS: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 76
  General Impression Marking Scale. 76
  Analytic Scales 76
  Narrative Writing: GIM Results 78
  Narrative Writing: Analytic Scale Results. . . . 80
  Process Writing: GIM Results 84
  Process Writing: Analytic Scale Results. 86

WRITTEN PRODUCTS: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 90
  Attribute Marking Scales. 90
  Descriptive Scales. 90
  Error Counts. 91
  Narrative Writing: Attribute and Descriptive Scales/Error Counts. 92
  Process Writing: Attribute and Descriptive Scales/Error Counts 96

SPELLING. 100
  Reporting of the 1982 Results 100

SENTENCE WRITING. 102
  Objectives at the Grade 9 Level. 102
  An Example 103
  Results for the Sentence Writing 103
    Sections A and B. 105
    Section C 106

WRITTEN PRODUCTS:

GENERAL IMPRESSION MARKING AND ANALYTIC SCALES

On the 1982 Manitoba Writing Assessment, students' written products were evaluated in a number of different ways in order to come to conclusions about the writing performance demonstrated by students. All student writing was evaluated on two types of scales: A General Impression Marking Scale (one score) and Analytic Marking Scales (nine scores). Thus, each student paper was assigned ten different scores for each written product (two each in grades 3, 6, and 9; one in grade 12).

The scales are described in detail in the Manual and Scoring Key at each grade level. A summary is given below.

General Impression Marking Scale (GIM)

General Impression Marking (GIM) is a judgement about the overall quality of a students' writing performance, given the purpose of the writing, the audience to which the writing was addressed, and the expectation for that performance at a grade level. On the provincial tests, two purposes were identified for students at grade levels 3, 6, and 9. The first purpose was to relate a story (narration); the second purpose was to describe or explain how to do something (process writing). At the grade 12 level, students were asked only to write to describe or explain a process. At all grade levels and on all tasks, the audience identified for the students was their class.

The GIM is a quick impressionistic score. Each written product from each student was read by two different raters quickly for a general reaction. No marks were placed on the papers themselves. Each rater assigned a score of 1 to 6 (1 - low, 6 - high). If the scores of the two raters did not differ more than one point (e.g., one rater assigned the paper a 5, the other assigned it a 4), the scores were added together (5 + 4 = 9). That summed score is the GIM score reported for students on the writing task. If the scores of the two raters did differ more than one point (e.g., one rater assigned the paper a 2, the other a 5), the paper was read by an arbitrator. (Arbitrators were scorers chosen because of their experience in doing this type of marking.) In the arbitrated cases, the GIM score for a piece of writing was calculated as follows:

$$\frac{\text{FIRST SCORE + SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{GIM SCORE}$$

The lowest score a written product could receive was 2 and the highest was 12.

The GIM score attempts to rate compositions in a general, global manner (See Cooper & Odell, 1977 for a complete discussion). Covered by the GIM are any qualities that impress the rater including the quality of thought evidenced in the paper, the effectiveness of the overall organization, the effectiveness of the overall presentation, and the effectiveness of the style and word choice.

Analytic Scales

Like general impression marking, the Analytic Scales are quick, qualitative judgements about the effectiveness of a piece of writing. Analytic Scales, however, are measures of the effectiveness of particular qualities of a paper. For the provincial writing assessment, nine different qualities were

selected for testing, five rhetorical or discourse categories, four mechanics and convention categories. These scales were developed because they seemed to be the qualities most often evaluated in writing by teachers and because they were objectives most often identified as being taught by teachers.

Like the GIM, a paper was assigned a score between 1 and 6 (1 - low, 6 - high), but for each of the nine qualities.

Rhetorical Categories

1. Organization and Sequence
2. Wording and Style
3. Focus or Theme
4. Ideas
5. Choice of Details and Elaboration of Ideas

Mechanics and Conventions Categories

6. Sentence Structure
7. Usage
8. Punctuation and Capitalization
9. Spelling

A scorer, usually different from the GIM scorers, read each written product for the Analytic Scales. Specific descriptions of the qualities associated with high, middle, and low papers are given in each grade level Manual and Scoring Key, as well as examples of papers. (For a more complete discussion of analytic scales in general, refer to Diederich, 1974, and Cooper & Odell, 1977.)

For the provincial sample, every paper was rated on all nine analytic scales by at least one rater. A subset of papers at each grade level was marked analytically by a second rater in order to calculate the amount of agreement or reliability between independent raters.

These two scoring procedures (GIM and Analytic) yielded ten different scores for each written product. At the grade 3, 6, and 9 levels, students wrote two pieces of extended writing so that each student in those grades received 20 different scores (10 for each writing product). At the grade 12 level, the students were asked to write one piece of extended writing so that each student received 10 different scores. This is summarized in Table 2 (p. 78).

The GIM and Analytic score results at the grade 9 level are based on scoring the total provincial sample of 1197 test papers.

TABLE 2

GIM AND ANALYTIC SCORING OF WRITTEN PRODUCTS

Scales	Narrative Writing				Descriptive/Expository Writing			
	3	6	9	12	3	6	9	12
GIM (Scores range from 2 to 12)	x	x	x	-	x	x	x	x
Analytic (Scores range from 1 to 6)								
1. Organization and sequence	x	x	x	-	x	x	x	x
2. Wording and style (Word Choice)	x	x	x	-	x	x	x	x
3. Focus or theme	x	x	x	-	x	x	x	x
4. Ideas	x	x	x	-	x	x	x	x
5. Choice of Details & Elaboration of Ideas	x	x	x	-	x	x	x	x
6. Sentence Structure	x	x	x	-	x	x	x	x
7. Usage	x	x	x	-	x	x	x	x
8. Punctuation & Capitalization	x	x	x	-	x	x	x	x
9. Spelling	x	x	x	-	x	x	x	x

Narrative Writing: GIM Results

Students at the grade 9 level were asked to write a story using the three-sentence story beginning provided as a stimulus (Part II of the test). Teachers were instructed to read the directions aloud to their students. Students were given approximately 50 minutes to write their stories.

The results of the GIM scoring for the story are presented on page 79. Scores for the two scorers combined ranged from 2 to 12, with scores 2, 3, and 4 described as "Low"; 5, 6, 7, 8, and 9 described as "Middle"; and 10, 11, and 12 described as "High".

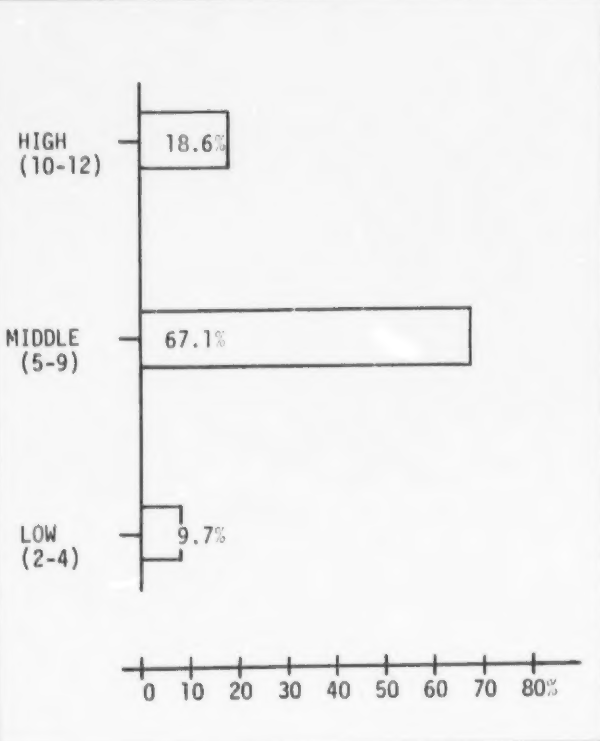
The results are presented in three ways:

1. The percentage of students obtaining each GIM score: e.g., 3.3% of students received a combined GIM score of 12 (both scorers awarded a 6 to these stories).
2. The percentage of students obtaining GIM scores categorized as High, Middle, and Low: e.g., 18.6% of students received combined GIM scores of 10, 11, or 12 (High).
3. The average (mean) performance of students on the GIM scale (and the standard deviation). These are based on the scale of 2 to 12.

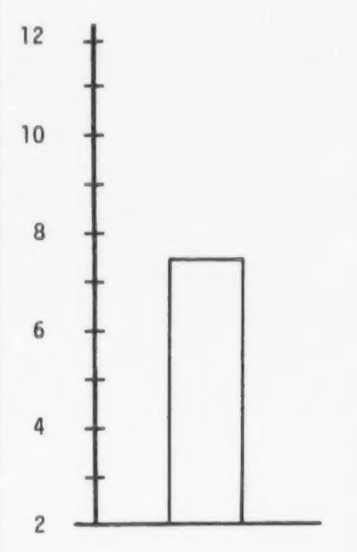
NARRATIVE WRITING: 1982 GIM RESULTS (Grade 9)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
HIGH	12	3.3%
	11	5.8
	10	9.5
MIDDLE	9	12.8
	8	12.1
	7	19.4
	6	12.5
	5	10.3
LOW	4	5.4
	3	2.5
	2	1.8
Insufficient Information		4.4

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 7.40
Standard Deviation = 2.29

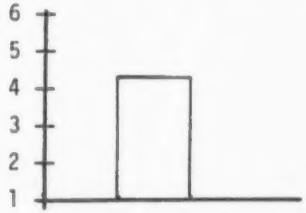
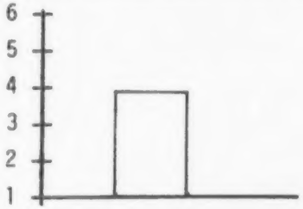
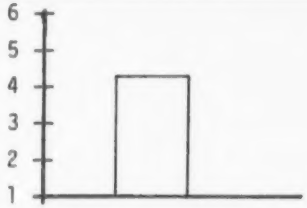
Story Writing: Analytic Scale Results

The results of the analytic scoring for the grade 9 story are presented on pages 81 to 83. All of the scores range on a scale from 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High".

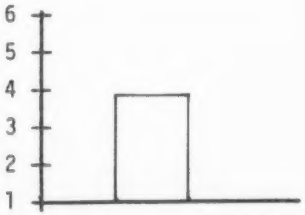
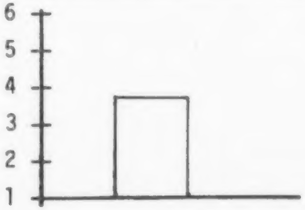
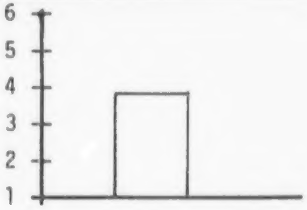
Results are presented in two ways:

1. Percentage of students obtaining each score: e.g., 10.1% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students, and standard deviation, based on the scale of 1 to 6: e.g., the average of students was 4.16 (standard deviation was 1.10) on Organization and Sequence.

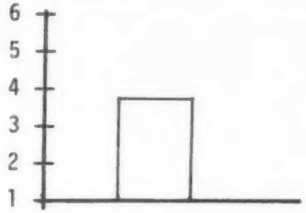
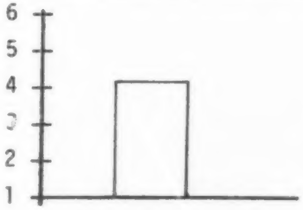
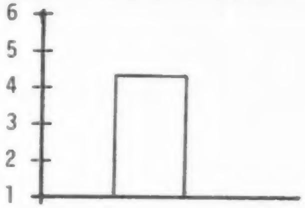
NARRATIVE WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 9)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	10.1%	
		5	30.1	
	MIDDLE	4	32.1	
		3	19.6	
	LOW	2	6.7	
		1	0.7	
	Insufficient Information		0.8	Average = 4.16 Standard Deviation = 1.10
2. Wording and Style	HIGH	6	7.7	
		5	24.2	
	MIDDLE	4	30.7	
		3	28.1	
	LOW	2	8.0	
		1	0.7	
	Insufficient Information		0.7	Average = 3.93 Standard Deviation = 1.10
3. Focus or Theme	HIGH	6	11.3	
		5	31.3	
	MIDDLE	4	30.1	
		3	17.9	
	LOW	2	6.8	
		1	1.3	
	Insufficient Information		1.4	Average = 4.19 Standard Deviation = 1.15

NARRATIVE WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 9)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	7.1	
		5	21.6	
	MIDDLE	4	33.1	
		3	27.4	
	LOW	2	7.4	
		1	0.9	
	Insufficient Information		2.6	Average = 3.91 Standard Deviation = 1.08
5. Choice of Details and Elaboration of Ideas	HIGH	6	6.6	
		5	21.3	
	MIDDLE	4	32.9	
		3	25.9	
	LOW	2	9.7	
		1	1.7	
	Insufficient Information		1.9	Average = 3.84 Standard Deviation = 1.13
6. Sentence Structure	HIGH	6	7.4	
		5	26.6	
	MIDDLE	4	26.9	
		3	25.1	
	LOW	2	11.5	
		1	1.8	
	Insufficient Information		0.8	Average = 3.88 Standard Deviation = 1.20

NARRATIVE WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 9)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	8.8%	
		5	26.3	
	MIDDLE	4	25.9	
		3	21.3	
	LOW	2	12.0	
		1	3.1	
	Insufficient Information		2.6	Average = 3.89 Standard Deviation = 1.27
8. Punctuation and Capitalization	HIGH	6	12.8	
		5	32.5	
	MIDDLE	4	21.6	
		3	18.1	
	LOW	2	11.2	
		1	3.0	
	Insufficient Information		0.8	Average = 4.09 Standard Deviation = 1.32
9. Spelling	HIGH	6	18.7	
		5	30.1	
	MIDDLE	4	19.9	
		3	14.2	
	LOW	2	9.2	
		1	3.7	
	Insufficient Information		4.2	Average = 4.25 Standard Deviation = 1.38

Process Writing: GIM Results

Students at the grade 9 level were asked to write directions about how to do something, using as suggested topics: How to shop for ______, How to make a ______, How to cook ______, or Getting ready for ______ (Part IV of the test). Teachers were instructed to read the directions aloud to their students. Students were given approximately 30 minutes to write their directions.

These student compositions were scored using the same GIM analytic scales as were used for the stories (see pages 76-78).

The results of the GIM scoring are presented on page 85. Scores for the two scorers combined ranged from 2 to 12, with scores 2, 3, and 4 categorized as "Low"; 5, 6, 7, 8 and 9 as "Middle"; and 10, 11, and 12 as "High".

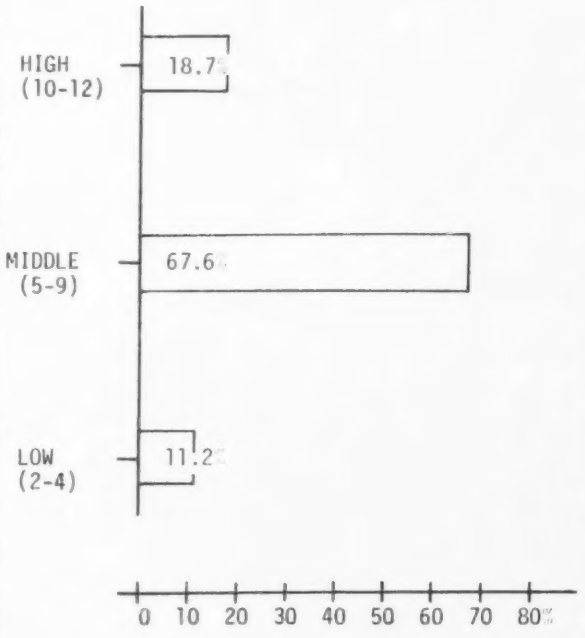
The results are depicted in three ways:

1. The percentage of students obtaining each GIM score: e.g., 2.3% of students received a combined GIM score of 12.
2. The percentage of students obtaining GIM scores categorized as High, Medium, and Low: e.g., 18.7% of students received scores in the High category (GIM score of 10, 11, or 12).
3. The average (mean) performance of students on the GIM scale (and the standard deviation).

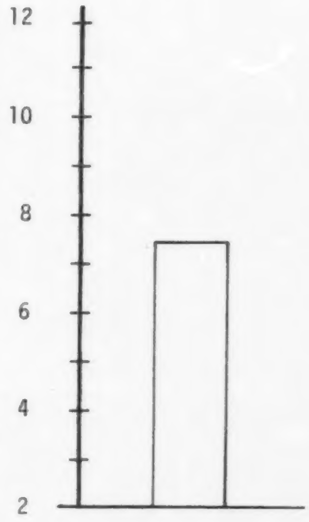
PROCESS WRITING: 1982 GIM RESULTS (Grade 9)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
HIGH	12	2.3%
	11	7.4
	10	9.0
MIDDLE	9	16.7
	8	11.4
	7	19.1
	6	10.2
	5	10.2
LOW	4	4.5
	3	4.4
	2	2.3
Insufficient Information		2.7

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 7.41

Standard Deviation = 2.35

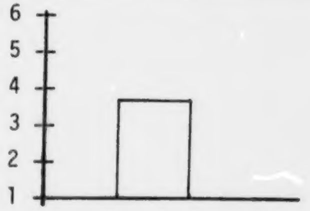
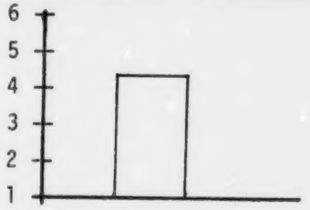
Process Writing: Analytic Scale Results

The results of the analytic scoring for the grade 9 process writing are presented on pages 87 to 89. All of the analytic scores are on a scale of 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High". These scales are identical to those used in evaluating the narrative writing (see pages 76 to 78 for a detailed description).

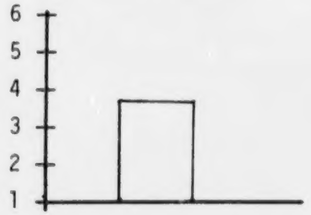
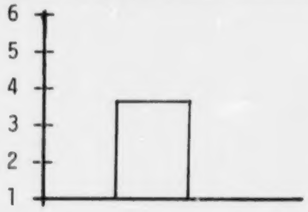
Results for each analytic scale are depicted in two ways:

1. Percentage of students obtaining each score: e.g., 8.2% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students on each analytic scale (and standard deviation) based on the scale from 1 to 6: e.g., the average performance was 3.95 (standard deviation was 1.23) on Organization and Sequence.

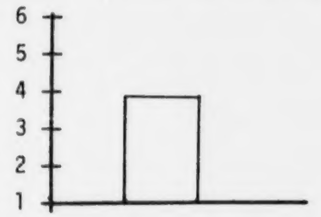
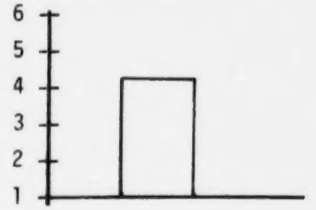
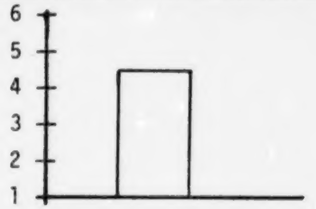
PROCESS WRITING: 1982 ANALYTIC SCALE RESULTS (Grade 9)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	8.2%	
		5	28.6	
	MIDDLE	4	28.4	
		3	21.4	
	LOW	2	9.1	
		1	3.4	
	Insufficient Information		0.9	Average = 3.95 Standard Deviation = 1.23
2. Wording and Style	HIGH	6	5.1	
		5	20.2	
	MIDDLE	4	32.3	
		3	27.6	
	LOW	2	10.0	
		1	3.8	
	Insufficient Information		1.1	Average = 3.71 Standard Deviation = 1.16
3. Focus or Theme	HIGH	6	15.2	
		5	33.0	
	MIDDLE	4	25.3	
		3	16.1	
	LOW	2	6.2	
		1	2.8	
	Insufficient Information		1.3	Average = 4.27 Standard Deviation = 1.24

PROCESS WRITING: 1982 ANALYTIC SCALE RESULTS

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	6.8	
		5	22.7	
	MIDDLE	4	31.1	
		3	24.0	
	LOW	2	10.3	
		1	3.4	
	Insufficient Information		1.7	Average = 3.81 Standard Deviation = 1.20
5. Choice of Details and Elaboration of Ideas	HIGH	6	5.2	
		5	19.9	
	MIDDLE	4	28.3	
		3	25.7	
	LOW	2	13.5	
		1	6.3	
	Insufficient Information		1.3	Average = 3.58 Standard Deviation = 1.27
6. Sentence Structure	HIGH	6	5.0	
		5	20.9	
	MIDDLE	4	27.8	
		3	27.3	
	LOW	2	14.5	
		1	3.4	
	Insufficient Information		1.2	Average = 3.64 Standard Deviation = 1.21

PROCESS WRITING: 1982 ANALYTIC SCALE RESULTS

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	6.8%	
		5	27.4	
	MIDDLE	4	29.1	
		3	23.1	
	LOW	2	9.3	
		1	2.8	
	Insufficient Information		1.5	Average = 3.91 Standard Deviation = 1.19
8. Punctuation and Capitalization	HIGH	6	14.8	
		5	28.8	
	MIDDLE	4	25.5	
		3	16.1	
	LOW	2	10.1	
		1	3.0	
	Insufficient Information		1.6	Average = 4.13 Standard Deviation = 1.32
9. Spelling	HIGH	6	21.7	
		5	33.0	
	MIDDLE	4	18.6	
		3	14.2	
	LOW	2	6.7	
		1	2.8	
	Insufficient Information		2.9	Average = 4.42 Standard Deviation = 1.32

WRITTEN PRODUCTS:

ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS

In addition to the qualitative judgements made on each paper using the GIM and Analytic Scales, a randomly selected sub-sample of papers was examined for more intensive structural, linguistic, attribute, and error descriptions. These in-depth analyses were carried out by the same group of raters employed for the other marking tasks. A total of 95 student papers were selected at the grade 9 level for this intensive scoring.

The rationale for these types of analyses was to make assessments regarding the kinds of structures students used in their actual writing and the kinds of errors most frequently made by students in actual writing situations. Also, they were conducted in order to draw some conclusions about the number of students demonstrating serious problems in the structural aspects of their writing. The choice of structures, errors, and attributes was based on the survey of writing done by Hunt (See Cooper & Odell, 1977), on the work of Evans (1979) in the Ontario assessment, and on the objectives identified by practicing teachers as the ones of greatest interest and use to them in their classrooms.

Attribute Marking Scales

Seven different attribute marking scales were identified as being of interest to teachers. The scales themselves were evaluations about whether or not a student was experiencing serious (in need of remedial help) difficulty in particular areas. For each paper evaluated, the rater was asked to rate the composition on the seven attributes, each on a three-point scale:

3 - No errors;

2 - An occasional error, but not rated as a serious problem;

1 - Serious problem (Student is in need of special remedial help);

X - Insufficient information (Structures not used or insufficient writing for an evaluation to be made).

Five of the seven areas were evaluated at all grade levels. The last two (use of contractions and use of quotation marks in direct discourse) were evaluated only at grade levels 3 and 6. The seven attributes were:

1. Run-on sentences
2. Use of complete sentences
3. Paragraphing
4. Internal punctuation
5. Capitalization
6. Use of contractions
7. Use of quotation marks in direct discourse

Descriptive Scales

The descriptive scales were counts of particular structures used by students. The following counts were made for each student composition in grades 3, 6, and 9. Grade 12 compositions were tallied on all but the last one.

1. Total number of words written
2. Total number of T-units (independent clauses) written
3. Total number of dependent clauses written
4. Total number of adverbial clauses written

5. Total number of relative clauses written
6. Total number of noun clauses written
7. Total number of modifiers used (not counted)
8. Total number of instances of direct discourse (not counted)

The particular counts were chosen because they seem closely related both to composition quality and to syntactic fluency (See Hunt in Cooper & Odell, 1977). The purpose was to judge whether or not students easily and fluently used these structures in their actual writing.

From these counts, a number of transformed scores were calculated to produce twelve indices of writing maturity in students. In general, the more fluent, mature writers will demonstrate higher scores on each of these indices, whereas weaker, less mature writers will score lower. These scores should be viewed, however, only as indicators of how students handle the syntax and grammar of the language in free writing situations. The question answered by these data is: Do students use these particular structures in their own writing? The question of whether or not students _can_ use these structures is left to be answered by the sentence writing tasks (Part III).

The scores calculated included:

Words per T-unit
Words per clause
Clauses per T-unit
Dependent clauses per T-unit
Adverbial clauses per T-unit
Adverbial clauses per 100 words
Relative clauses per T-unit
Relative clauses per 100 words
Noun clauses per T-unit
Noun clauses per 100 words
Direct discourse per T-unit
Direct discourse per 100 words

Error Counts

Error counts, though tedious to do, can give information about how well students avoid error in actual writing situations. These counts are realistic alternatives to the use of objective (particularly multiple-choice type) items in evaluating writing. Error counts can give greater information about the choices that students are making in actual writing tasks.

Each of the error categories are calculated in these ways:

average number of errors per composition
average number of errors per T-unit (i.e. per independent clause), and
average number of errors per 100 running words

The error counts obtained were:

1. Number of sentence fragments
2. Number of incorrectly placed modifiers
3. Number of sentences not beginning with a capital letter
4. Number of proper nouns not capitalized
5. Number of sentences not ending with appropriate punctuation
6. Number of internal comma errors
7. Number of faulty subject-verb agreement
8. Number of verb tense errors
9. Number of word order errors

Similarly to the descriptive scales a number of scores were calculated from the error counts. But, opposite to the descriptive scales, higher scores on the scale reflect greater error rates.

The calculated scores were:

Sentence fragments per T-unit
Sentence fragments per 100 words
Incorrectly placed modifiers per T-unit
Incorrectly placed modifiers per 100 words
Beginning capitalization errors per T-unit
Beginning capitalization errors per 100 words
Proper nouns not capitalized per T-unit
Proper nouns not capitalized per 100 words
End punctuation errors per T-unit
End punctuation errors per 100 words
Comma errors per T-unit
Comma errors per 100 words
Faulty agreement per T-unit
Faulty agreement per 100 words
Verb tense errors per T-unit
Verb tense errors per 100 words
Word order errors per T-unit
Word order errors per 100 words

Narrative Writing: Attribute and Descriptive Scales/Error Counts

The results for the grade 9 stories (Part II) are presented for:

The 7 attribute scales (p. 93),

The 8 descriptive scales (p. 94),

The 12 indices calculated from the descriptive scales (p. 94),

The 9 error counts (p. 95), and

The 18 indices calculated from the error counts (p. 95).

Since there are a large number of scales, the results have been tabulated in a compacted form in the following tables.

The average (mean) and standard deviation are reported for each scale and index. For the attribute scales, the percentage of students obtaining each score is reported: e.g., 32.6% of students obtained a score of 3 (no errors) on the attribute scale "Run-on sentences", and the average is 2.04 on the scale of 1 to 3 (standard deviation is 0.80).

NARRATIVE WRITING: 1982 ATTRIBUTE SCALES RESULTS (Grade 9)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
1. Run-on sentences				
No errors	3	32.6		
An occasional error	2	35.8	2.04	0.80
Serious problem	1	28.4		
Insufficient information		3.2		
2. Use of complete sentences				
No errors	3	52.6		
An occasional error	2	34.7	2.46	0.66
Serious problem	1	8.4		
Insufficient information		4.2		
3. Paragraphing				
No errors	3	52.6		
An occasional error	2	28.4	2.40	0.74
Serious problem	1	14.7		
Insufficient information		4.2		
4. Internal punctuation				
No errors	3	13.7		
An occasional error	2	61.1	1.93	0.60
Serious problem	1	20.0		
Insufficient information		5.3		
5. Capitalization				
No errors	3	54.7		
An occasional error	2	33.7	2.48	0.65
Serious problem	1	8.4		
Insufficient information		3.2		
6. Use of contractions*				
No errors	3	15.8		
An occasional error	2	2.1	-	-
Serious problem	1	1.1		
Insufficient information		81.0		
7. Use of quotation marks in direct discourse*				
No errors	3	-		
An occasional error	2	-	-	-
Serious problem	1	-		
Insufficient information		-		

*Many students did not use these particular features in their writing; this was categorized as "Insufficient information".

NARRATIVE WRITING: DESCRIPTIVE SCALE RESULTS (Grade 9)

Descriptive Scale (Actual counts)	Average Number Mean	Standard Deviation
1. Words	355.22	129.26
2. T-units	35.20	14.44
3. Dependent Clauses	15.94	11.33
4. Adverbial clauses	6.94	4.62
5. Relative clauses	1.65	1.76
6. Noun clauses	5.63	4.95
7. Modifiers	(not counted)	
8. Instances of direct discourse	(not counted)	

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	10.53	2.88
Words per clause	7.22	1.89
Clauses per T-unit	1.48	0.32
Dependent clauses per T-unit	0.48	0.32
Adverbial clauses per T-unit	0.22	0.18
Adverbial clauses per 100 words	2.07	1.69
Relative clauses per T-unit	0.053	0.058
Relative clauses per 100 words	0.48	0.47
Noun clauses per T-unit	0.16	0.12
Noun clauses per 100 words	1.54	1.09
Direct discourse per T-unit	-	-
Direct discourse per 100 words	-	-

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

NARRATIVE WRITING: ERROR COUNTS (Grade 9)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.016	0.038
Sentence fragments per 100 words	0.21	0.74
Incorrectly placed modifiers per T-unit	0.013	0.032
Incorrectly placed modifiers per 100 words	0.11	0.26
Beginning capitalization errors per T-unit	0.088	0.119
Beginning capitalization errors per 100 words	0.85	1.18
Proper nouns not capitalized per T-unit	0.009	0.075
Proper nouns not capitalized per 100 words	0.068	0.565
End punctuation errors per T-unit	0.096	0.111
End punctuation errors per 100 words	0.93	1.08
Comma errors per T-unit	0.15	0.16
Comma errors per 100 words	1.40	1.57
Faulty agreement per T-unit	0.014	0.031
Faulty agreement per 100 words	0.13	0.31
Verb tense errors per T-unit	0.023	0.050
Verb tense errors per 100 words	0.22	0.46
Word order errors per T-unit	0.012	0.038
Word order errors per 100 words	0.095	0.251

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Average Error Count Standard Deviation
1. Sentence fragments	0	73.9%	0.60	1.58
	1	13.0		
	2 or more	13.1		
2. Incorrectly placed modifiers	0	77.2	0.42	0.92
	1	9.8		
	2 or more	13.0		
3. Sentences not beginning with Capital letters	0	41.3	3.21	4.84
	1	9.8		
	2 or more	48.9		
4. Proper nouns not capitalized	0	95.7	0.22	1.88
	1	3.3		
	2 or more	1.1		
5. Sentences not ending with appropriate punctuation	0	33.7	3.50	4.69
	1	10.9		
	2 or more	55.4		
6. Internal comma errors	0	17.4	4.45	4.78
	1	6.5		
	2 or more	76.1		
7. Instances of faulty subject/verb agreement	0	76.1	0.52	1.20
	1	10.9		
	2 or more	13.0		
8. Verb tense errors	0	65.2	0.80	1.60
	1	18.5		
	2 or more	16.3		
9. Word order errors	0	80.4	0.35	0.91
	1	12.0		
	2 or more	7.6		

Process Writing: Attribute and Descriptive Scales/Error Counts

The results for the process writing part of the grade 9 test (Part IV) are presented for the scales and calculated indices. The scales are identical to those used for the stories (for a description see pages 90 to 92):

The 7 attribute scales (p. 97),
The 8 descriptive scales (p. 98),
The 12 indices calculated from the descriptive scales (p. 98),
The 9 error counts (p. 99), and
The 18 indices calculated from the error counts (p. 99).

Due to the large number of scales and indices the results have been tabulated in a compacted form in the following tables. The average (mean) and standard deviation are reported for each scale and index. As well, for the attribute scales, the percentage of students obtaining each score is reported: e.g., 35.8% of students obtained a score of 3 (no errors) on the attribute scale "Run-on Sentences", and the average is 2.19 on the scale of 1 to 3 (standard deviation 0.74).

PROCESS WRITING: 1982 ATTRIBUTE SCALES RESULTS (Grade 9)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
1. Run-on sentences				
No errors	3	35.8%		
An occasional error	2	40.0	2.19	0.74
Serious problem	1	17.9		
Insufficient information		6.3		
2. Use of complete sentences				
No errors	3	45.3		
An occasional error	2	41.1	2.40	0.64
Serious problem	1	7.4		
Insufficient information		6.3		
3. Paragraphing				
No errors	3	51.6		
An occasional error	2	28.4	2.44	0.75
Serious problem	1	11.6		
Insufficient information		8.5		
4. Internal punctuation				
No errors	3	21.1		
An occasional error	2	61.1	2.11	0.58
Serious problem	1	10.5		
Insufficient information		7.4		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance	
			Mean	Standard Deviation
5. Capitalization				
No errors	3	58.9%		
An occasional error	2	31.6	2.60	0.56
Serious problem	1	3.2		
Insufficient information		6.3		
6. Use of contractions*				
No errors	3	11.6		
An occasional error	2	5.3	-	-
Serious problem	1	0.0		
Insufficient information		83.2		
7. Use of quotation marks in direct discourse*				
No errors	3	10.5		
An occasional error	2	1.1	-	-
Serious problem	1	0.0		
Insufficient information		88.4		

*Many students did not use these particular features in their writing; this was categorized as "Insufficient information".

PROCESS WRITING: 1982 DESCRIPTIVE SCALE RESULTS

	Descriptive Scale (Actual counts)	Average Number	
		Mean	Standard Deviation
1.	Words	183.48	87.26
2.	T-units	15.07	8.59
3.	Dependent Clauses	8.64	6.44
4.	Adverbial clauses	4.60	3.02
5.	Relative clauses	1.58	1.77
6.	Noun clauses	2.06	3.33
7.	Modifiers	(not counted)	
8.	Instances of direct discourse	(not counted)	

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	13.01	3.81
Words per clause	8.14	1.78
Clauses per T-unit	1.63	0.46
Dependent clauses per T-unit	0.63	0.46
Adverbial clauses per T-unit	0.35	0.28
Adverbial clauses per 100 words	2.53	1.57
Relative clauses per T-unit	0.12	0.16
Relative clauses per 100 words	0.84	0.98
Noun clauses per T-unit	0.14	0.18
Noun clauses per 100 words	1.09	1.59
Direct discourse per T-unit	-	-
Direct discourse per 100 words	-	-

*"Clause" includes independent and dependent clauses; i.e., Number of T-units + Number of dependent clauses (Descriptive Scale 2 + Descriptive Scale 3).

PROCESS WRITING: 1982 ERROR COUNTS

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.031	0.057
Sentence fragments per 100 words	0.22	0.40
Incorrectly placed modifiers per T-unit	0.018	0.051
Incorrectly placed modifiers per 100 words	0.14	0.35
Beginning capitalization errors per T-unit	0.10	0.17
Beginning capitalization errors per 100 words	0.86	1.58
Proper nouns not capitalized per T-unit	0.004	0.023
Proper nouns not capitalized per 100 words	0.030	0.149
End punctuation errors per T-unit	0.12	0.18
End punctuation errors per 100 words	1.08	1.59
Comma errors per T-unit	0.24	0.21
Comma errors per 100 words	1.82	1.62
Faulty agreement per T-unit	0.015	0.051
Faulty agreement per 100 words	0.11	0.34
Verb tense errors per T-unit	0.022	0.056
Verb tense errors per 100 words	0.16	0.43
Word order errors per T-unit	0.032	0.074
Word order errors per 100 words	0.23	0.50

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Average Error Count Standard Deviation
1. Sentence fragments	0	69.2%	0.46	0.92
	1	23.1		
	2 or more	7.7		
2. Incorrectly placed modifiers	0	83.5	0.29	0.76
	1	9.9		
	2 or more	6.6		
3. Sentences not beginning with Capital letters	0	51.6	1.46	2.69
	1	17.6		
	2 or more	30.8		
4. Proper nouns not capitalized	0	94.5	0.08	0.34
	1	3.3		
	2 or more	2.2		
5. Sentences not ending with appropriate punctuation	0	44.0	1.88	2.89
	1	17.6		
	2 or more	38.4		
6. Internal comma errors	0	22.0	2.97	2.75
	1	8.8		
	2 or more	69.2		
7. Instances of faulty subject/verb agreement	0	84.6	0.22	0.63
	1	12.1		
	2 or more	3.3		
8. Verb tense errors	0	81.3	0.41	1.03
	1	9.9		
	2 or more	8.8		
9. Word order errors	0	76.9	0.45	1.13
	1	15.4		
	2 or more	7.7		

SPELLING

The purpose of the spelling assessment on the 1982 Writing Assessment Program was to determine if students could spell correctly the words they actually used in their writing as well as the words they were likely to need in their writing. Spelling in the actual written products was assessed by one of the analytic scales (Analytic Scales, p. 76). In order to assess the performance of students on the spelling of words they were likely to need in their general writing (but may not have used in the writing tasks provided on the test), words were randomly selected from Ves Thomas' Canadian Word Lists (1979). This particular set of word lists assigns words to different grade levels based on the frequency of the appearance of these words in the writing of students at a particular grade level. For example, the grade 9 list is composed of words that a high number of Canadian grade 9 students use in their own writing.

In addition to the words selected from the Thomas Canadian Word Lists, a number of words were selected from the 1979 Manitoba Writing Assessment Program so that comparisons between student performance in 1979 and 1982 could be made. These 1979 words were originally selected not because they should be in students' *writing* vocabularies but because they conformed to particular patterns and appeared in grade level *reading* materials: that is, they were words that students could be expected to recognize and pronounce (but not necessarily be able to use or spell in writing). The rationale for the selection of the 1982 spelling words seemed stronger in that it reflected words that students were likely to use in composing rather than in reading.

The spelling test used in the 1982 assessment contained two separate tests (Part I of the test). One set of items (those marked with a * in the *Manual and Scoring Key* was taken from the 1979 Manitoba Writing Assessment. The total number of these words is 6.

The balance of the words were randomly selected from the Thomas list at the grade 9 level. There were 20 words obtained in this way for the 1982 component.

Two scores were given on the test, one for the items from 1979 (Total possible = 6), another for the items selected for 1982 (Total possible = 20). Total overall performance (Total possible = 26) was determined by adding the two scores together. The results are based on the provincial sample of 1197 grade 9 students.

Reporting of the 1982 Results

The average (mean) score is reported for the scores based on the 6 items, on the 20 items, and on the 26 items (combination of 6 and 20 items). A frequency distribution indicating the percentage of students obtaining each score is presented (e.g., 24.0% of students spelled 23 or 24 of the words correctly). Finally, the percentage of students spelling each word correctly is indicated in a separate table on page 101.

SUMMARY OF THE 1982 SPELLING RESULTS (Grade 9)

Score Correct out of Total Number of Words (26)	Percentage of Students Obtaining Each Score
25-26	30.9%
23-24	24.0
21-22	15.1
19-20	10.2
17-18	6.1
15-16	4.4
13-14	3.6
11-12	2.3
9-10	1.5
7-8	0.6
5-6	0.8
3-4	0.1
1-2	0.5
0	0.2

Average Score Correct	21.5
Standard Deviation	4.65
Average Percentage Correct	82.8

Words Selected From the 1979 Assessment		Words Selected For the 1982 Assessment	
Items	Percentage of 1982 students obtaining item correct	Items	Percentage of 1982 students obtaining item correct
5. tragedy	63.0	1. local	98.0
8. pamphlet	64.5	2. complaining	91.0
12. provincial	79.2	3. current	90.2
16. innocent	73.7	4. foreign	63.7
19. parliament	52.0	6. mountains	93.6
22. dessert	63.4	7. require	89.7
		9. value	93.1
		10. camera	92.9
		11. easily	91.6
		13. personality	91.3
		14. solar	96.5
		15. allowance	81.7
		17. citizens	81.2
		18. confusion	84.0
		20. deserve	90.3
		21. original	83.0
		23. private	90.2
		24. scene	83.9
		25. supermarkets	90.9
		26. weighed	79.2

Average Correct	3.96	Average Correct	17.56
Standard Deviation	1.75	Standard Deviation	3.25
Average Percentage Correct	66.0%	Average Percentage Correct	87.8%

SENTENCE WRITING

A number of sentence writing tasks were designed to assess performance on selected structural, syntactic, and usage objectives. The particular format of the sentence writing tasks was chosen for five reasons:

> It closely resembled the kind of thinking and shaping a writer probably goes through when composing
>
> It allowed for a variety of answers which could be labeled "correct" so that a student was not limited to a single correct answer and could create the most effective sentence(s) for the situation and information provided.
>
> The structurally correct answers in most cases could be generally categorized from least mature (or fluent) to most mature (or fluent).
>
> The open format centred on what students could do in particular circumstances rather than on what they could not do.
>
> The format could allow for the assessment of whether students could use particular structures.

Performance on the sentence writing tasks could be compared with the descriptive scales in the actual writing tasks in order to determine if students in fact did use the targeted structures in free writing.

Although a number of objectives (or targeted structures) were common across all four grade levels, the particular structures or conventions assessed at each grade level were based on objectives that practicing teachers felt were applicable at that level.

Objectives at the Grade 9 Level

Objectives at the grade 9 level are listed below; the sections comprised Part III of the student test.

Section A - Objective: Use of participles
Section B - Objective: Use of relativization
Section C - Objective 1: Use of nominal coordination
Section C - Objective 2: Use of adverbial subordination

Several sentences were required for each section but some sentences were scored for two objectives.

The objectives fall into one major area: syntactic maturity. All students should have some understanding of, as well as considerable skill in, this area.

Students' sentences were scored according to several categories; these were denoted a, b, c, etc. It is likely that the most careful, mature writers would respond consistently with answers that would fall into category "a" of each objective. Less mature writers might have a profile of mixed performance (e.g., demonstrate sentence sense and knowledge of punctuation, but not consistently demonstrate the ability to capitalize correctly or use the relatively mature structures such as adverbial subordination). The fact that a student did not consistently answer with the most mature structures does not necessarily mean that she/he does not know how to employ the structure or convention, rather that the structure or convention is not readily accessible in the written mode. The very weakest students might demonstrate little fluency in the structures they write or choose. The tasks presented in the assessment allow for divergent answers so that the teacher can assess the level of maturity or fluency of the student(s).

An Example

For Item 1 of Section C of the grade 9 test the student was asked to combine the following two sentences using one of the words suggested.

> 1. Jerry will help you with the painting.
> Carla will not help you with the painting.

The most fluent writer would probably have responded with an answer such as:

> Jerry, but not Carla, will help you with the painting.

This answer would be categorized a: Combined using nominal coordination.

Another writer, however, might have responded with:

> Although Carla will not help you with the painting, Jerry will.

By no means is this a less mature answer than the one that was categorized a; however, it is sufficiently different that it would seem useless to categorize the two structures together in a simple "correct" category. This answer, because it employed adverbial subordination, would be categorized as b: Combined using subordination.

A somewhat less fluent writer might have responded with an answer such as:

> Jerry will help you with the painting but
> Carla will not help you with the painting.

This answer would be categorized c: Combined using clausal coordination. This answer is still structurally correct and grammatical; it is simply not as efficient or fluent as either of the first two answers. The categorization scheme for this objective, then, allows the teacher to judge a structure as "correct", but "less mature".

Yet another writer might have answered:

> Jerry's helping you with the painting means that Carla will not.

Here, the writer has clearly not followed directions, but the purpose of the writing assessment is not to evaluate direction-following behavior, but writing behavior. The student has written a structurally correct, grammatical, and meaningful sentence, and should be given credit for it. This answer, in fact, might be thought of as more mature than any previously cited. This answer would be categorized as d: Correctly combined in some other fashion. The maturity or fluency level is not apparent from this category description, but the scheme does give the writer credit for meaningful writing.

The final category, e: Incorrectly combined, was only used if the student wrote an answer that did not make sense or was not grammatical.

The categories for scoring varied according to the particular objective, but in all cases the scheme was similar to the example presented. The complete categories and examples of answers are given in the grade 9 Manual and Scoring Key.

Results for the Sentence Writing

At the grade 9 level students were asked to do two or more sentence writing tasks (items) for each of the objectives. The results are presented for each sentence and for each objective. Since the scoring categories for each objective were not based on a numerical scale averages cannot be used. Instead reported are the percentages of students whose sentences were categorized in each of the scoring categories (the first scoring category is always the targeted structure of the objective). Also reported is the average of the percentages for a particular category for the two items pertaining to an objective. For example, in Section A 81.7% of student sentences were categorized as "a: Correctly combined using participles" for Item 1 and 76.0% for Item 2. On average, 78.9% of student sentences were categorized "a: Correctly combined using participles" for the two sentences combined, and thus for the objective: Use of participles.

Results are based on the total sample of 1197 test papers.

SENTENCE WRITING: 1982 RESULTS

(Grade 9)

Scale categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section A-Objective: Use of participles			
	1. The boys carried their cameras. The boys were excited. The cameras were hanging around their necks.	2. The students cheered the team to victory. The students were yelling at the top of their lungs. The team was exhausted.	
a - Correctly combined using participles	81.7%	76.0%	78.9%
b - Correctly combined in some other fashion	12.4	15.5	13.9
c - Incorrectly combined	4.5	6.6	5.6
Omitted	1.5	1.9	1.7
Section B-Objective: Use of relativization			
	1. The ambulance stopped in front of the house. The ambulance's siren was wailing. The house was burning.	2. The pilot planned a route. The pilot had just arrived at the hanger. The route would take him across the border.	
a - Correctly combined using relativization	73.0%	80.2%	76.6%
b - Correctly combined in some other fashion	6.3	2.1	4.2
c - Incorrectly combined	18.3	14.5	16.4
Omitted	2.5	3.3	2.9

SENTENCE WRITING: 1982 RESULTS (Grade 9)

Scale Categories	Percentage of students obtaining each category for the Items			Percentage of students obtaining each category: Average percentage for the items
Section C-Objective: Nominal coordination				
	1. Jerry will help you with the painting. Carla will not help you with the painting.	3. I will pick up some ice cream for the party. I will not pick up the cake for the party.		
a - Combined using nominal coordination	21.5%	8.9%		15.2%
b - Combined using subordination	39.7	11.9		25.8
c - Combined using clausal coordination	35.9	74.5		55.2
d - Combined in some other fashion	0.8	1.0		0.9
e - Incorrectly combined	0.5	0.7		0.6
Omitted	1.6	3.0		2.3
Section C-Objective: Use of adverbial subordination				
	2. We had to finish the painting last night. It looked like rain.	4. The oak tree was struck by lightning. The oak tree fell onto the old house. We moved the furniture out of the house.	5. The movie was late in starting out. The crowd became nervous. The crowd began to shuffle their feet. The crowd began to throw popcorn.	
a - Adverbial subordination used	92.4%	82.0%	72.4%	82.3%
b - Coordination used	4.4	9.3	13.9	9.2
c - Combined in some other fashion	0.6	2.3	3.2	2.0
d - Incorrectly combined	0.6	4.1	5.8	3.5
Omitted	2.0	2.3	4.7	3.0

CHAPTER 5

GRADE 12 TEST: RESULTS FOR ENGLISH 300 STUDENTS

The grade 12 writing test consisted of four parts:

Part I - Spelling: list of 20 spelling words dictated by the teacher.

Part II - Essay Planning: planning of discourse including title, theme, supporting details, and concluding paragraph based on provided information.

Part III - Sentence Writing: sentence production based on lists of words and on short sentences.

Part IV - Essay Writing: written product (process writing) explaining how to do something.

Parts II and IV resulted in written products. The product for Part II, essay planning, was scored according to three criteria, whereas for Part IV, the essay was scored on the same scales as were the compositions at the other grade levels.

Results of the 1982 assessment are presented in the order specified below. The results are provided in separate chapters for students in English 300 and 301. A brief description of the scales and scoring procedures precedes the presentation of results.

WRITTEN PRODUCT: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES. 108
  General Impression Marking Scale (GIM) 108
  Analytic Scales. 108
  Essay Writing: GIM results. 110
  Essay Writing: Analytic Scale Results 112
WRITTEN PRODUCT: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 116
  Attribute Marking Scales 116
  Descriptive Scales 116
  Error Counts 117
  Essay Writing: Attribute and Descriptive Scales/Error Counts. 118
ESSAY PLANNING 122
  Essay Planning: Results. 122
SPELLING 126
  Reporting of the 1982 Results 126
SENTENCE WRITING 128
  Objectives at the Grade 12 Level. 128
  An Example. 129
  Results for the Sentence Writing. 129
    Section A. 131
    Section B. 131
    Section C. 132
    Section D. 132

WRITTEN PRODUCTS:

GENERAL IMPRESSION MARKING AND ANALYTIC SCALES

On the 1982 Manitoba Writing Assessment, students' written products were evaluated in a number of different ways in order to come to conclusions about the writing performance demonstrated by students. All student writing was evaluated on two types of scales: A General Impression Marking Scale (one score) and Analytic Marking Scales (nine scores). Thus, each student paper was assigned ten different scores for each written product (two each in grades 3, 6, and 9; one in grade 12).

The scales are described in detail in the Manual and Scoring Key at each grade level. A summary is given below.

General Impression Marking Scale (GIM)

General Impression Marking (GIM) is a judgement about the overall quality of a students' writing performance, given the purpose of the writing, the audience to which the writing was addressed, and the expectation for that performance at a grade level. On the provincial tests, two purposes were identified for students at grade levels 3, 6, and 9. The first purpose was to relate a story (narration); the second purpose was to describe or explain how to do something (process writing). At the grade 12 level, students were asked only to write to describe or explain a process. At all grade levels and on all tasks, the audience identified for the students was their class.

The GIM is a quick impressionistic score. Each written product from each student was read by two different raters quickly for a general reaction. No marks were placed on the papers themselves. Each rater assigned a score of 1 to 6 (1 - low, 6 - high). If the scores of the two raters did not differ more than one point (e.g., one rater assigned the paper a 5 the other assigned it a 4), the scores were added together (5 + 4 = 9). That summed score is the GIM score reported for students on the writing task. If the scores of the two raters did differ more than one point (e.g., one rater assigned the paper a 2, the other a 5), the paper was read by an arbitrator. (Arbitrators were scorers chosen because of their experience in doing this type of marking.) In the arbitrated cases, the GIM score for a piece of writing was calculated as follows:

$$\frac{\text{FIRST SCORE} + \text{SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{GIM SCORE}$$

The lowest score a written product could receive was 2 and the highest was 12.

The GIM score attempts to rate compositions in a general, global manner (See Cooper & Odell, 1977 for a complete discussion). Covered by the GIM are any qualities that impress the rater including the quality of thought evidenced in the paper, the effectiveness of the overall organization, the effectiveness of the overall presentation, and the effectiveness of the style and word choice.

Analytic Scales

Like General Impression Marking, the Analytic Scales are quick, qualitative judgements about the effectiveness of a piece of writing. Analytic Scales, however, are measures of the effectiveness of particular qualities of a paper. For the provincial writing assessment, nine different qualities were

selected for testing, five rhetorical or discourse categories, four mechanics and convention categories. These scales were developed because they seemed to be the qualities most often evaluated in writing by teachers and because they were objectives most often identified as being taught by teachers.

Like the GIM, a paper was assigned a score between 1 and 6 (1 - low, 6 - high), but for each of the nine qualities.

Rhetorical Categories

1. Organization and Sequence
2. Wording and Style
3. Focus or Theme
4. Ideas
5. Choice of Details and Elaboration of Ideas

Mechanics and Conventions Categories

6. Sentence Structure
7. Usage
8. Punctuation and Capitalization
9. Spelling

A scorer, usually different from the GIM scorers, read each written product for the Analytic Scales. Specific descriptions of the qualities associated with high, middle, and low papers are given in each grade level Manual and Scoring Key, as well as examples of papers. (For a more complete discussion of analytic scales in general, refer to Diederich, 1974, and Cooper & Odell, 1977.)

For the provincial sample, every paper was rated on all nine analytic scales by at least one rater. A subset of papers at each grade level was marked analytically by a second rater in order to calculate the amount of agreement or reliability between independent raters.

These two scoring procedures (GIM and Analytic) yielded ten different scores for each written product. At the grade 3, 6, and 9 levels, students wrote two pieces of extended writing so that each student in those grades received 20 different scores (10 for each writing product). At the grade 12 level, the students were asked to write one piece of extended writing so that each student received 10 different scores. This is summarized in Table 2 (p. 110).

The GIM and Analytic score results for students in English 300 are based on scoring the total English 300 sample of 451 test papers.

TABLE 2

GIM AND ANALYTIC SCORING OF WRITTEN PRODUCTS

Scales	Narrative Writing				Descriptive/Expository Writing			
	3	6	9	12	3	6	9	12
GIM (Scores range from 2 to 12)	x	x	x	-	x	x	x	x
Analytic (Scores range from 1 to 6)								
1. Organization and sequence	x	x	x	-	x	x	x	x
2. Wording and style (Word Choice)	x	x	x	-	x	x	x	x
3. Focus or theme	x	x	x	-	x	x	x	x
4. Ideas	x	x	x	-	x	x	x	x
5. Choice of Details & Elaboration of Ideas	x	x	x	-	x	x	x	x
6. Sentence Structure	x	x	x	-	x	x	x	x
7. Usage	x	x	x	-	x	x	x	x
8. Punctuation & Capitalization	x	x	x	-	x	x	x	x
9. Spelling	x	x	x	-	x	x	x	x

Essay Writing: GIM Results

Students at the grade 12 level were asked to write an essay explaining how something is done (Part IV of the test). Topics were suggested, such as: How to shop for a used car, How to train for your favourite sport, How to help a person in distress, etc. The criteria used in the Analytic scales were also given. Students were given approximately 60 minutes to write their stories.

The results of the GIM scoring for the essay are presented on page 111. Scores for the two scorers combined ranged from 2 to 12, with scores, 2, 3, and 4 described as "Low"; 5, 6, 7, 8, and 9 described as "Middle"; and 10, 11, and 12 described as "High".

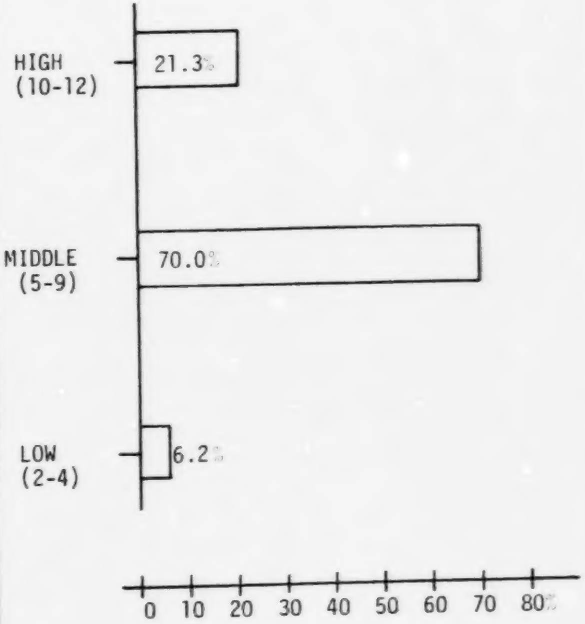
The results are presented in three ways:

1. The percentage of students obtaining each GIM score: e.g., 2.09% of English 300 students received a combined GIM score of 12 (both scorers awarded a 6 to these stories).
2. The percentage of students obtaining GIM scores categorized as High, Middle, and Low: e.g., 21.3% of students received combined GIM scores of 10, 11, or 12 (High).
3. The average (mean) performance of students on the GIM scale (and the standard deviation). These are based on the scale of 2 to 12.

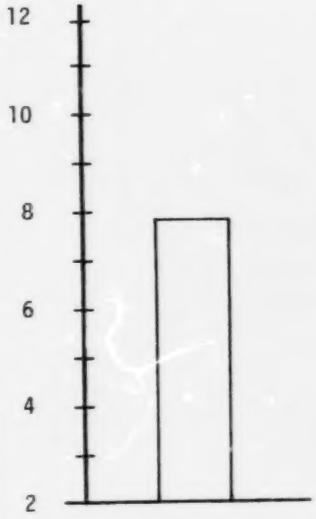
ESSAY WRITING: 1982 GIM RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

GIM Scale (two scorers (combined)		Percentage of 1982 Students
		Obtaining Each Score
	12	2.0%
HIGH	11	7.8
	10	11.5
	9	20.4
	8	17.5
MIDDLE	7	16.1
	6	8.0
	5	8.0
	4	4.2
LOW	3	1.8
	2	0.2
Insufficient Information		2.5

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 7.92

Standard Deviation = 2.04

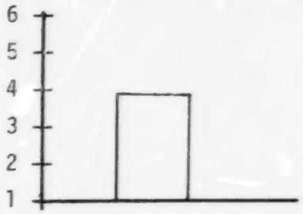
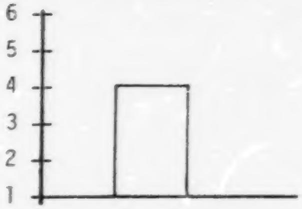
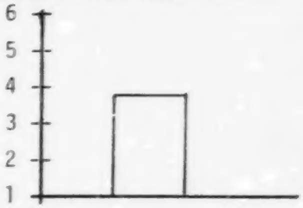
Essay Writing: Analytic Scale Results

The results of the analytic scoring for the grade 12 essay are presented on pages 113 to 115. All of the scores range on a scale from 1 to 6. Scores 1 and 2 are described as "low", 3 and 4 as "middle", and 5 and 6 as "high".

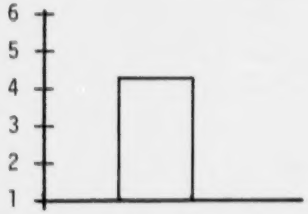
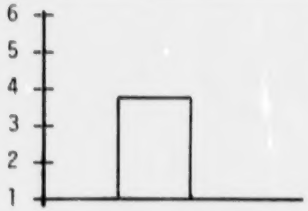
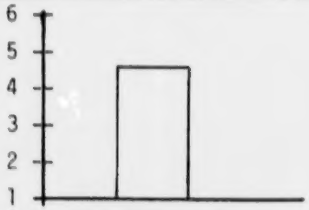
Results are presented in two ways:

1. Percentage of students obtaining each score: e.g., 9.3% of students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students, and standard deviation, based on the scale of 1 to 6: e.g., the average of students was 4.19 (standard deviation was 1.06) on Organization and Sequence.

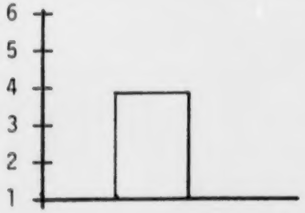
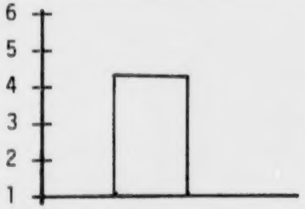
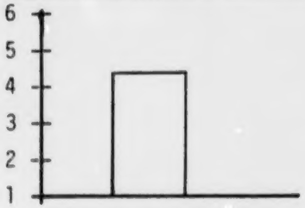
ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	6.2	
		5	20.6	
	MIDDLE	4	40.6	
		3	25.6	
	LOW	2	4.4	
		1	0.2	
	Insufficient Information		1.3	Average = 3.97 Standard Deviation = 0.97
5. Choice of Details and Elaboration of Ideas	HIGH	6	7.1	
		5	23.9	
	MIDDLE	4	37.9	
		3	23.7	
	LOW	2	4.2	
		1	1.3	
	Insufficient Information		1.8	Average = 4.02 Standard Deviation = 1.03
6. Sentence Structure	HIGH	6	3.3	
		5	24.4	
	MIDDLE	4	35.9	
		3	25.9	
	LOW	2	7.8	
		1	1.6	
	Insufficient Information		1.1	Average = 3.85 Standard Deviation = 1.04

ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	9.3	
		5	29.9	
	MIDDLE	4	35.3	
		3	18.6	
	LOW	2	4.2	
		1	1.1	
	Insufficient Information		1.6	Average = 4.19 Standard Deviation = 1.06
2. Wording and Style	HIGH	6	3.1	
		5	19.7	
	MIDDLE	4	41.2	
		3	27.5	
	LOW	2	6.7	
		1	0.4	
	Insufficient Information		1.3	Average = 3.84 Standard Deviation = 0.94
3. Focus or Theme	HIGH	6	15.5	
		5	43.2	
	MIDDLE	4	26.2	
		3	10.6	
	LOW	2	2.7	
		1	0.2	
	Insufficient Information		1.5	Average = 4.59 Standard Deviation = 0.98

ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	4.9	
		5	28.2	
	MIDDLE	4	31.3	
		3	25.1	
	LOW	2	8.0	
		1	1.3	
	Insufficient Information		1.4	Average = 3.93 Standard Deviation = 1.09
8. Punctuation and Capitalization	HIGH	6	9.3	
		5	37.5	
	MIDDLE	4	29.9	
		3	17.3	
	LOW	2	3.8	
		1	1.1	
	Insufficient Information		1.1	Average = 4.28 Standard Deviation = 1.05
9. Spelling	HIGH	6	15.5	
		5	34.8	
	MIDDLE	4	26.8	
		3	14.4	
	LOW	2	4.2	
		1	2.4	
	Insufficient Information		1.8	Average = 4.36 Standard Deviation = 1.18

WRITTEN PRODUCT:

ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS

In addition to the qualitative judgements made on each paper using the GIM and Analytic Scales, a randomly selected sub-sample of papers was examined for more intensive structural, linguistic, attribute, and error descriptions. These in-depth analyses were carried out by the same group of raters employed for the other marking tasks. A total of 55 student papers were selected at the grade 12 level for this intensive scoring.

The rationale for these types of analyses was to make assessments regarding the kinds of structures students used in their actual writing and the kinds of errors most frequently made by students in actual writing situations. Also, they were conducted in order to draw some conclusions about the number of students demonstrating serious problems in the structural aspects of their writing. The choice of structures, errors, and attributes was based on the survey of writing done by Hunt (See Cooper & Odell, 1977), on the work of Evans (1979) in the Ontario assessment, and on the objectives identified by practicing teachers as the ones of greatest interest and use to them in their classrooms.

Attribute Marking Scales

Seven different attribute marking scales were identified as being of interest to teachers. The scales themselves were evaluations about whether or not a student was experiencing serious (in need of remedial help) difficulty in particular areas. For each paper evaluated, the rater was asked to rate the composition on the seven attributes, each on a three-point scale:

3 - No errors;
2 - An occasional error, but not rated as a serious problem
1 - Serious problem (Student is in need of special remedial help);
X - Insufficient information (Structures not used or insufficient writing for an evaluation to me made).

Five of the seven areas awere evaluated at all grade levels. The last two (use of contractions and use of quotation marks in direct discourse) were evaluated only at grade levels 3 and 6. The seven attributes were:

1. Run-on sentences
2. Use of complete sentences
3. Paragraphing
4. Internal punctuation
5. Capitalization
6. Use of contractions
7. Use of quotation marks in direct discourse

Descriptive Scales

The descriptive scales were counts of particular structures used by students. The following counts were made for each student composition in grades 3, 6, and 9. Grade 12 compositions were tallied on all but the last one.

1. Total number of words written
2. Total number of T-units (independent clauses) written
3. Total number of dependent clauses written
4. Total number of adverbial clauses written

5. Total number of relative clauses written
6. Total number of noun clauses written
7. Total number of modifiers used (not counted)
8. Total number of instances of direct discourse (not counted)

The particular counts were chosen because they seem closely related both to composition quality and to syntactic fluency (See Hunt in Cooper & Odell, 1977). The purpose was to judge whether or not students easily and fluently used these structures in their actual writing.

From these counts, a number of transformed scores were calculated to produce twelve indices of writing maturity in students. In general, the more fluent, mature writers will demonstrate higher scores on each of these indices, whereas weaker, less mature writers will score lower. These scores should be viewed, however, only as indicators of how students handle the syntax and grammar of the language in free writing situations. The question answered by these data is: Do students use these particular structures in their own writing? The question of whether or not students can use these structures is left to be answered by the sentence writing tasks (Part III).

The scores calculated included:

Words per T-unit
Words per clause
Clauses per T-unit
Dependent clauses per T-unit
Adverbial clauses per T-unit
Adverbial clauses per 100 words
Relative clauses per T-unit
Relative clauses per 100 words
Noun clauses per T-unit
Noun clauses per 100 words
Direct discourse per T-unit
Direct discourse per 100 words

Error Counts

Error counts, though tedious to do, can give information about how well students avoid error in actual writing situations. These counts are realistic alternatives to the use of objective (particularly multiple-choice type) items in evaluating writing. Error counts can give greater information about the choices that students are making in actual writing tasks.

Each of the error categories are calculated in these ways:

average number of errors per composition
average number of errors per T-unit (i.e. per independent clause), and
average number of errors per 100 running words

The error counts obtained were:

1. Number of sentence fragments
2. Number of incorrectly placed modifiers
3. Number of sentences not beginning with a capital letter
4. Number of proper nouns not capitalized
5. Number of sentences not ending with appropriate punctuation
6. Number of internal comma errors
7. Number of faulty subject-verb agreement
8. Number of verb tense errors
9. Number of word order errors

Similarly to the descriptive scales a number of scores were calculated from the error counts. But, opposite to the descriptive scales, higher scores on the scale reflect greater error rates.

The calculated scores were:

Sentence fragments per T-unit
Sentence fragments per 100 words
Incorrectly placed modifiers per T-unit
Incorrectly placed modifiers per 100 words
Beginning capitalization errors per T-unit
Beginning capitalization errors per 100 words
Proper nouns not capitalized per T-unit
Proper nouns not capitalized per 100 words
End punctuation errors per T-unit
End punctuation errors per 100 words
Comma errors per T-unit
Comma errors per 100 words
Faulty agreement per T-unit
Faulty agreement per 100 words
Verb tense errors per T-unit
Verb tense errors per 100 words
Word order errors per T-unit
Word order errors per 100 words

Essay Writing: Attribute and Descriptive Scales/Error Counts

The results for the essays of the English 300 students (Part IV) are presented for:

The 7 attribute scales (p. 119),
The 8 descriptive scales (p. 120),
The 12 indices calculated from the descriptive scales (p. 120),
The 9 error counts (p. 121), and
The 18 indices calculated from the error counts (p. 121).

Since there are a large number of scales, the results have been tabulated in a compacted form in the following tables.

The average (mean) and standard deviation are reported for each scale and index. For the attribute scales, the percentage of students obtaining each score is reported: e.g., 51.6% of students obtained a score of 3 (no errors) on the attribute scale "Run-on sentences", and the average is 2.52 on the scale of 1 to 3 (standard deviation is 0.51.

Of the 55 essays intensively scored at the grade 12 level, 31 are from English 300 students. The results are based on these 31 essays.

ESSAY WRITING: 1982 ATTRIBUTE SCALES RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
1. Run-on sentences				
No errors	3	51.6%		
An occasional error	2	48.4	2.52	0.51
Serious problem	1	0.0		
Insufficient information		0.0		
2. Use of complete sentences				
No errors	3	71.0		
An occasional error	2	22.6	2.70	0.54
Serious problem	1	3.2		
Insufficient information		3.2		
3. Paragraphing				
No errors	3	74.2		
An occasional error	2	22.6	2.71	0.53
Serious problem	1	3.2		
Insufficient information		0.0		
4. Internal punctuation				
No errors	3	9.7		
An occasional error	2	74.2	1.94	0.51
Serious problem	1	16.1		
Insufficient information		0.0		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
5. Capitalization				
No errors	3	67.7%		
An occasional error	2	29.0	2.70	0.47
Serious problem	1	0.0		
Insufficient information		3.2		
6. Use of contractions				
No errors	3			
An occasional error	2			
Serious problem	1		not counted	
Insufficient information				
7. Use of quotation marks in direct discourse				
No errors	3			
An occasional error	2		not counted	
Serious problem	1			
Insufficient information				

ESSAY WRITING: DESCRIPTIVE SCALE RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Descriptive Scale (Actual counts)	Average Number	
	Mean	Standard Deviation
1. Words	520.68	119.71
2. T-units	37.61	11.56
3. Dependent Clauses	23.07	9.66
4. Adverbial clauses	11.16	5.05
5. Relative clauses	4.84	3.00
6. Noun clauses	4.39	3.79
7. Modifiers	(not counted)	
8. Instances of direct discourse	(not counted)	

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	14.58	3.94
Words per clause	8.76	1.66
Clauses per T-unit	1.68	0.43
Dependent clauses per T-unit	0.68	0.43
Adverbial clauses per T-unit	0.32	0.16
Adverbial clauses per 100 words	2.25	1.13
Relative clauses per T-unit	0.138	0.092
Relative clauses per 100 words	0.93	0.55
Noun clauses per T-unit	0.12	0.13
Noun clauses per 100 words	0.82	0.68
Direct discourse per T-unit	-	-
Direct discourse per 100 words	-	-

ESSAY WRITING: ERROR COUNTS FOR ENGLISH 300 STUDENTS (Grade 12)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.009	0.021
Sentence fragments per 100 words	0.062	0.157
Incorrectly placed modifiers per T-unit	0.010	0.027
Incorrectly placed modifiers per 100 words	0.076	0.210
Beginning capitalization errors per T-unit	0.017	0.025
Beginning capitalization errors per 100 words	0.13	0.19
Proper nouns not capitalized per T-unit	0.0	-
Proper nouns not capitalized per 100 words	0.0	-
End punctuation errors per T-unit	0.015	0.025
End punctuation errors per 100 words	0.11	0.18
Comma errors per T-unit	0.16	0.14
Comma errors per 100 words	1.13	0.96
Faulty agreement per T-unit	0.011	0.019
Faulty agreement per 100 words	0.090	0.154
Verb tense errors per T-unit	0.006	0.015
Verb tense errors per 100 words	0.046	0.110
Word order errors per T-unit	0.013	0.050
Word order errors per 100 words	0.10	0.38

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Standard Deviation
1. Sentence fragments	0	83.9%	0.36	1.02
	1	6.5		
	2 or more	9.7		
2. Incorrectly placed modifiers	0	80.6	0.45	1.29
	1	12.9		
	2 or more	6.4		
3. Sentences not beginning with Capital letters	0	58.1	0.71	1.04
	1	22.6		
	2 or more	19.3		
4. Proper nouns not capitalized	0	100.0	0.00	-
	1	0.0		
	2 or more	0.0		
5. Sentences not ending with appropriate punctuation	0	64.5	0.65	1.14
	1	19.4		
	2 or more	16.1		
6. Internal comma errors	0	6.5	6.13	5.58
	1	12.9		
	2 or more	80.6		
7. Instances of faulty subject/ verb agreement	0	71.0	0.45	0.77
	1	12.9		
	2 or more	16.1		
8. Verb tense errors	0	83.9	0.26	0.63
	1	6.5		
	2 or more	9.7		
9. Word order errors	0	90.3	0.61	2.32
	1	0.0		
	2 or more	9.6		

ESSAY PLANNING

The purpose of the discourse planning task at the grade 12 level was to assess students' ability in pre-planning an essay. Since pre-writing and planning is an important aspect of the act of composing, it was decided that students at grade 12 could be asked to plan a persuasive essay. The topic was whether students in high school should hold part-time jobs.

In the task itself, students were given a list of ideas for the essay. Both important and unimportant ideas on both sides of the issue were interspersed. Students were then asked to plan an essay by identifying a title, a major theme, and a number of major ideas supporting that theme, and then to group important details under each of the major ideas. They were also asked to write the concluding paragraph of the essay.

The essay plan was rated using three analytic marking scales; these were indicated to students:

A. Organization and sequence,
B. Focus or theme, and
C. Choice of details and elaboration.

Raters read the essay plans, then reacted to them holistically by assigning a score of 1 to 6 (1 - low, 6 - high) for each of the three qualities. Particular attributes of high, middle, and low plans are described in the grade 12 Manual and Scoring Key.

All essay plans were rated by two different scorers in much the same manner as was done for the general impression marking of the essay (see p. 108). A third scorer was assigned to arbitrate the essay plan if the first two scorers differed by two or more points on two or three of the scales.

The scores of the first two scorers were combined to form a scale that ranged from 2 (1 + 1) to 12 (6 + 6). In the arbitrated cases the score was calculated by:

$$\frac{\text{FIRST SCORE + SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{COMBINED SCORE}$$

Essay Planning: Results

Teachers were instructed to read the directions aloud to students for the essay planning task (Part II). Students were given approximately 45 minutes to complete their plan. These plans were rated on the three scales indicated above.

The results of the 451 English 300 papers are presented on pages 123 to 125 in these three ways:

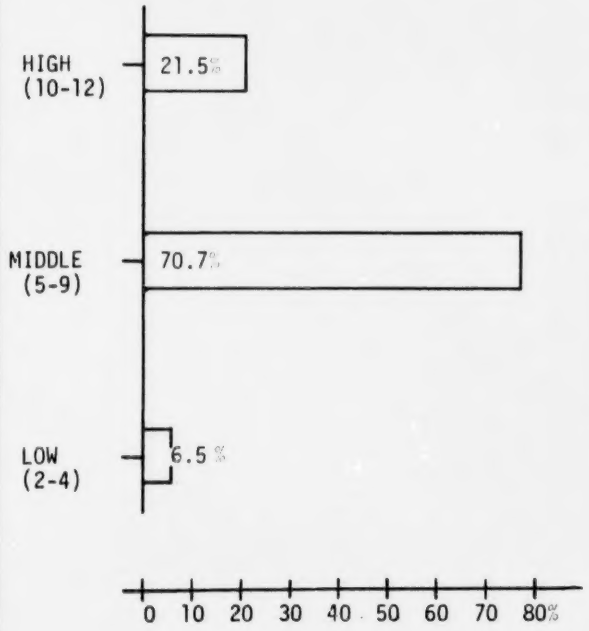
1. The percentage of students obtaining each combined score: e.g., on the scale "Organization and Sequence", 1.3% of English 300 students received a combined score of 12 (both scorers rated it a 6).
2. The percentage of students obtaining combined scores categorized as High (10-12), Middle (5-9), and Low (2-4): on "Organization and Sequence", 21.5% of English 300 students were in "High" category.
3. The average (mean) performance of students on each scale (and standard deviation). These are based on the scale of 2 to 12.

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

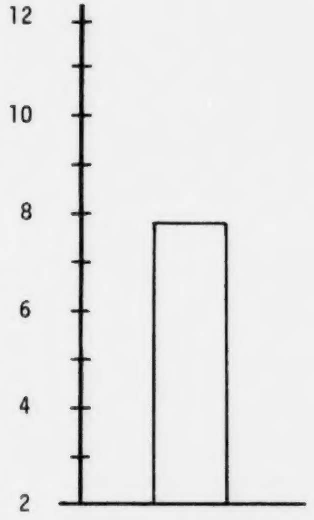
A. ORGANIZATION AND SEQUENCE

GIM Scale (two scorers (combined)		Percentage of 1982 Students
		Obtaining Each Score
	12	1.3%
HIGH	11	6.7
	10	13.5
	9	19.9
	8	17.5
MIDDLE	7	16.2
	6	9.8
	5	7.3
	4	3.1
LOW	3	2.7
	2	0.7
Insufficient Information		1.3

Obtaining Each Score Category



Average Performance



Average = 7.87

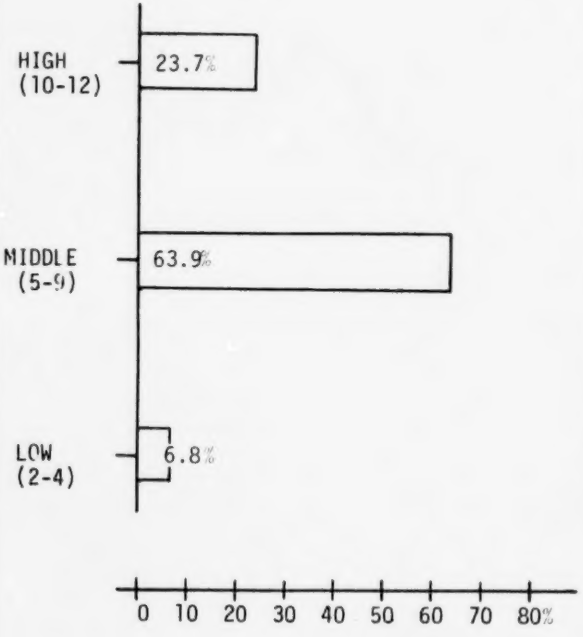
Standard Deviation = 2.03

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

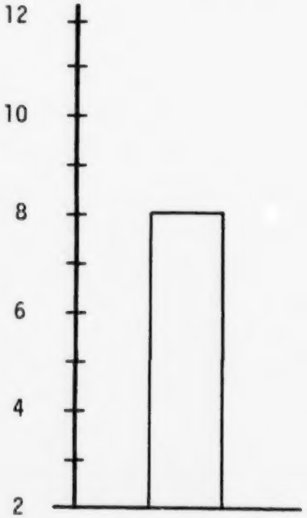
B. FOCUS OR THEME

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
	12	1.8%
HIGH	11	7.8
	10	14.1
	9	20.4
	8	18.4
MIDDLE	7	12.2
	6	6.7
	5	6.2
	4	3.9
LOW	3	2.0
	2	0.9
Insufficient Information		5.5

Percentage of 1982 Students Obtaining Each Score Category



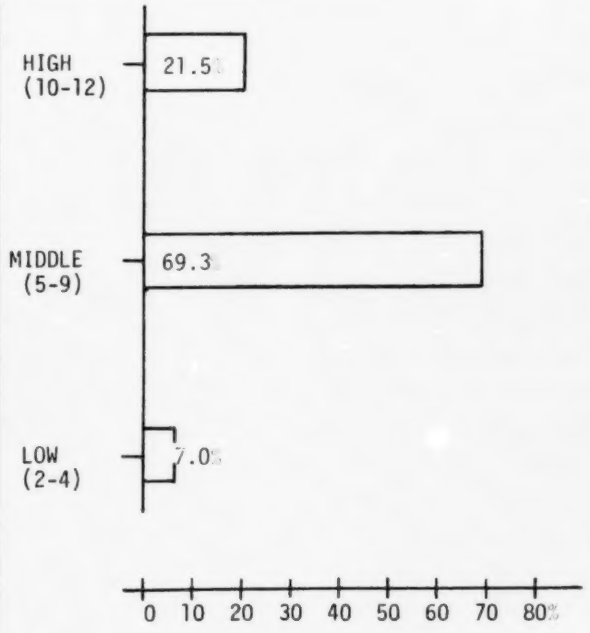
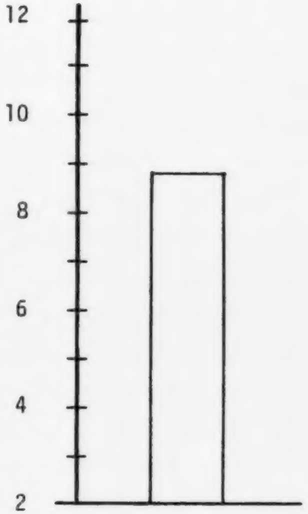
Average Performance



Average = 8.06
Standard Deviation = 2.10

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

C. CHOICE OF DETAIL AND ELABORATION

GIM Scale (two scorers (combined)		Percentage of 1982 Students: Obtaining Each Score	Percentage of 1982 Students: Obtaining Each Score Category	Average Performance
	12	2.2%		
HIGH	11	7.3		
	10	12.0		
	9	17.9		
	8	16.8		
MIDDLE	7	16.2		
	6	10.0		
	5	8.4		
	4	3.7		
LOW	3	2.2		
	2	1.1		
Insufficient Information		2.0		Average = 7.80 Standard Deviation = 2.14

SPELLING

The purpose of the spelling assessment on the 1982 Writing Assessment Program was to determine if students could spell correctly the words they actually used in their writing as well as the words they were likely to need in their writing. Spelling in the actual written product at the Grade 12 level was assessed by one of the nine analytic scales (See Analytic Scales). In order to assess the performance of students on the spelling of words they were likely to need in their general writing (but may not have used in the writing tasks provided on the test), words were randomly selected from the vocabulary lists by Dale and O'Rourke (1976). This list of words assigns words to different grade levels based on vocabulary recognition tests generally given in North America. It was chosen because no recent survey of words that Grade 12 students use in general writing was available, either for Canada or the United States.

Unlike the 1979 assessment, the 1982 Manitoba Writing Assessment Program tests spelling by dictation (rather than by error recognition and correction in a multiple-choice format). This decision was based on the conclusions regarding the most valid form of testing spelling drawn by Thomas (1979). Unfortunately, because of the difference in test format, realistic comparisons between 1979 spelling performance and 1982 spelling performance at the Grade 12 level cannot be made. The words chosen for the 1982 assessment, however, seem more reflective of the words students need in general writing than those selected for the 1979 assessment.

There is no breakdown of the spelling words into categories. Thus, the results are based on the number of correctly spelled words out of the total of 20.

Reporting of the 1982 Results

The average (mean) score is reported for the scores based on the 20 words. A frequency distribution indicating the percentage of students obtaining each score is presented (e.g., 17.9% of English 300 students spelled 19 or 20 of the words correctly). Finally, the percentage of students spelling each word correctly is indicated in a separate table.

The results for the 451 English 300 students are given on page 127.

SUMMARY OF THE 1982 SPELLING RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Score Correct out of Total Number of Words (20)	Percentage of Students Obtaining Each Score
19-20	17.9%
17-18	24.8
15-16	23.0
13-14	16.9
11-12	9.5
9-10	4.4
7-8	1.5
5-6	1.1
3-4	0.4
1-2	0.0
0	0.2

Words Selected For the 1982 Assessment	
Items	Percentage of 1982 students obtaining item correct
1. basically	92.2%
2. possessed	72.5
3. prescriptive	87.8
4. wheelbase	98.4
5. urchin	63.4
6. bolster	87.8
7. winsome	57.0
8. nimble	90.7
9. deluge	74.1
10. nectarine	67.6
11. incapacitate	62.1
12. ungrudging	89.4
13. decrepit	47.7
14. croon	70.1
15. collective	98.2
16. slipshod	70.7
17. sedimentary	81.2
18. boycott	72.9
19. entreat	90.5
20. itinerary	67.8

Average Score Correct	15.4
Standard Deviation	3.30
Average Percentage Correct	77.1%

SENTENCE WRITING

A number of sentence writing tasks were designed to assess performance on selected structural, syntactic, and usage objectives. The particular format of the sentence writing tasks was chosen for five reasons:

> It closely resembled the kind of thinking and shaping a writer probably goes through when composing.
>
> It allowed for a variety of answers which could be labeled "correct" so that a student was not limited to a single correct answer and could create the most effective sentence(s) for the situation and information provided.
>
> The structurally correct answers in most cases could be generally categorized from least mature (or fluent) to most mature (or fluent).
>
> The open format centred on what students could do in particular circumstances rather than on what they could not do.
>
> The format could allow for the assessment of whether students could use particular structures.

Performance on the sentence writing tasks could be compared with the descriptive scales in the actual writing tasks in order to determine if students in fact did use the targeted structures in free writing.

Although a number of objectives (or targeted structures) were common across all four grade levels, the particular structures or conventions assessed at each grade level were based on objectives that practicing teachers felt were applicable at that level.

Objectives at the Grade 12 Level

Objectives at the grade 12 level are listed below; the sections comprised Part III of the student test.

Section A - Objective: Use of participles
Section B - Objective: Use of parallel structure
Section C - Objective: Use of relativization
Section D - Objective: Use of adverbial subordination

Several sentences were required for each section to provide a more accurate indicator of student ability.

The objectives fall into one major area: syntactic maturity. All students should have some understanding of, as well as considerable skill in, this area.

Students' sentences were scored according to several categories; these were denoted a, b, c, etc. It is likely that the most careful, mature writers would respond consistently with answers that would fall into category "a" of each objective. Less mature writers might have a profile of mixed performance (e.g., use participles correctly and form sentences with adverbial subordination but not consistently use parallel structure or relativization). The fact that a student did not consistently answer with the most mature structures does not necessarily mean that she/he does not know how to employ the structure or convention, rather that the structure or convention is not readily accessible in the written mode. The very weakest students might demonstrate little fluency in the structures they write or choose. The tasks presented in the assessment allow for divergent answers so that the teacher can assess the level of maturity or fluency of students.

An Example

For Item 7 of Section D of the grade 12 test the student was asked to combine the following two sentences using one of the words suggested.

7. There was an excited and cheerful atmosphere in the gym. The home team had won the basketball game.

The most fluent writer would probably have responded with an answer such as:

> There was an excited and cheerful atmosphere in the gym because the home team had won the basketball game.

This answer would be categorized a: Combined using adverbial subordination.

A somewhat less fluent writer, however might have responded with an answer such as:

> There was an excited and cheerful atmosphere in the gym and the home team had won the basketball game.

This answer would be categorized b: Correctly combined using only coordination. This answer is structurally correct and grammatical; it is simply not as efficient or fluent as the first answer. The categorization scheme for this objective, then, allows the teacher to judge a structure as "correct", but "less mature".

Another writer might have responded with:

> The home team having won the basketball game, there was an excited and cheerful atmosphere in the gym.

By no means is this a less mature answer than the one categorized as "a"; however, it is different enough to warrent being categorized separately. Since it was impossible to anticipate all "correct" answers, a category, c: Correctly combined in some other fashion, was devised to cover the answers not covered by "a" or "b". Here the student has written a structurally correct, grammatical, and meaningful sentence and should be given credit for it, even though it did not reflect the objective that we as test writers had in mind. The maturity or fluency level is not apparent from the category description; the above answer might be categorized with a somewhat less mature answer. It is included in order to give the writer credit for meaningful writing.

The final category, d: Incorrectly combined, was used only if the student wrote an answer that did not make sense or was not grammatical.

The categories for scoring varied according to the particular objective, but in all cases the scheme was similar to the example presented. The complete categories and examples of answers are given in the grade 12 Manual and Scoring Key.

Results for the Sentence Writing

At the grade 12 level students were asked to do two or three writing tasks (items) for each objective. The results are presented for each sentence and for each objective. Since the scoring categories for each objective were not based on a numerical scale averages cannot be used. Instead reported are the percentages of students whose sentences were categorized in each of the scoring categories (the first scoring category is always the targeted structure of the objective). Also reported is the average of the percentages for a particular category for the two items pertaining to an objective. For example, in Section A 92.7% of English 300 student sentences were categorized as a: Used participles for Item 1 and 92.2% for Item 2. On average, 92.5% of student sentences were categorized a: Used participles for the two sentences combined, and thus for the objective: Use of participles.

Results are based on the total English 300 sample of 451 test papers.

SENTENCE WRITING: 1982 RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section A-Objective: Use of participles			
	1. The quarterback leaned back and threw the pass. The quarterback was standing in the clear. The pass was the winning one.	2. The hunters pushed the canoe out into the river and began to paddle toward the opposite shore. The hunters were keeping an eye out for deer. The river was rushing.	
a - Participles used	92.7%	92.2%	92.5%
b - Correctly combined in some other fashion	4.7	4.0	4.4
c - Incorrectly combined	2.4	3.3	2.9
Omitted	0.2	0.4	0.3
Section B-Objective: Use of parallel structure			
	3. Harold couldn't remember where he sat in class. Harold couldn't remember whom he sat beside in class. Harold couldn't even remember the classroom number.	4. Centuries ago people thought the world was flat. They also thought that the oceans were filled with huge monsters. People thought if they sailed too far away from land, they would fall off the world.	
a - Correctly combined with structure parallel	57.6%	57.4%	57.5%
b - Correctly combined in some fashion not requiring parallelism	10.0	16.6	13.3
c - Structures not parallel	25.1	17.1	21.1
d - Incorrectly combined	6.9	8.4	7.6
Omitted	0.4	0.4	0.4

SENTENCE WRITING: 1982 RESULTS FOR ENGLISH 300 STUDENTS (Grade 12)

Scale Categories	Percentage of students obtaining each category for the Items			Percentage of students obtaining each category: Average percentage for the items
Section C-Objective: Use of relativization				
	5. The young woman stood up and began to sing a rousing song. The young woman's hair was bleached white. No one knew the song.	6. Four dark creatures scurried up past the old wall and began quickly to climb the rocks. The creatures looked both evil and scared. The rocks seemed to grow along a wide ditch.		
a - Relativization used	96.7%	96.0%		96.4%
b - Correctly combined in some other fashion	0.4	2.0		1.2
c - Incorrectly combined	2.4	1.8		2.1
Omitted	0.4	0.2		0.3
Section D-Objective: Adverbial subordination				
	7. There was an excited and cheerful atmosphere in the gym. The home team had won the basketball game.	8. I was feeling depressed. I had not studied. The teacher announced that the test would be postponed.	9. The halls were filled with papers. Students kept tripping over the trash. The class decided to begin a clean-up campaign.	
a - Adverbial subordination used	80.5%	79.4%	81.2%	80.4%
b - Correctly combined using only coordination	7.5	6.0	4.4	6.0
c - Correctly combined in some other fashion	8.0	5.5	6.2	6.7
d - Incorrectly combined	3.8	8.6	7.3	6.7
Omitted	0.2	0.4	0.8	4.7

CHAPTER 6

GRADE 12 TEST: RESULTS FOR ENGLISH 301 STUDENTS

The grade 12 writing test consisted of four parts:

Part I - Spelling: list of 20 spelling words dictated by the teacher.

Part II - Essay Planning: planning of discourse including title, theme, supporting details, and concluding paragraph based on provided information.

Part III - Sentence Writing: sentence production based on lists of words and on short sentences.

Part IV - Essay Writing: written product (process writing) explaining how to do something.

Parts II and IV resulted in written products. The product for Part II, essay planning, was scored according to three criteria, whereas for Part IV, the essay was scored on the same scales as were the compositions at the other grade levels.

Results of the 1982 assessment are presented in the order specified below. The results are provided in separate chapters for students in English 300 and 301. A brief description of the scales and scoring procedures precedes the presentation of results.

WRITTEN PRODUCT: GENERAL IMPRESSION MARKING AND ANALYTIC SCALES 134
  General Impression Marking Scale (GIM) 134
  Analytic Scales. 134
  Essay Writing: GIM results. 136
  Essay Writing: Analytic Scale Results 138
WRITTEN PRODUCT: ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS 142
  Attribute Marking Scales 142
  Descriptive Scales 142
  Error Counts 143
  Essay Writing: Attribute and Descriptive Scales/Error Counts. 144
ESSAY PLANNING 148
  Essay Planning: Results. 148
SPELLING 152
  Reporting of the 1982 Results 152
SENTENCE WRITING 154
  Objectives at the Grade 12 Level. 154
  An Example. 155
  Results for the Sentence Writing. 155
    Section A. 157
    Section B. 157
    Section C. 158
    Section D. 158

WRITTEN PRODUCTS:
GENERAL IMPRESSION MARKING AND ANALYTIC SCALES

On the 1982 Manitoba Writing Assessment, students' written products were evaluated in a number of different ways in order to come to conclusions about the writing performance demonstrated by students. All student writing was evaluated on two types of scales: A General Impression Marking Scale (one score) and Analytic Marking Scales (nine scores). Thus, each student paper was assigned ten different scores for each written product (two each in grades 3, 6, and 9; one in grade 12).

The scales are described in detail in the Manual and Scoring Key at each grade level. A summary is given below.

General Impression Marking Scale (GIM)

General Impression Marking (GIM) is a judgement about the overall quality of a students' writing performance, given the purpose of the writing, the audience to which the writing was addressed, and the expectation for that performance at a grade level. On the provincial tests, two purposes were identified for students at grade levels 3, 6, and 9. The first purpose was to relate a story (narration); the second purpose was to describe or explain how to do something (process writing). At the grade 12 level, students were asked only to write to describe or explain a process. At all grade levels and on all tasks, the audience identified for the students was their class.

The GIM is a quick impressionistic score. Each written product from each student was read by two different raters quickly for a general reaction. No marks were placed on the papers themselves. Each rater assigned a score of 1 to 6 (1 - low, 6 - high). If the scores of the two raters did not differ more than one point (e.g., one rater assigned the paper a 5 the other assigned it a 4), the scores were added together (5 + 4 = 9). That summed score is the GIM score reported for students on the writing task. If the scores of the two raters did differ more than one point (e.g., one rater assigned the paper a 2, the other a 5), the paper was read by an arbitrator. (Arbitrators were scorers chosen because of their experience in doing this type of marking.) In the arbitrated cases, the GIM score for a piece of writing was calculated as follows:

$$\frac{\text{FIRST SCORE + SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} = \text{GIM SCORE}$$

The lowest score a written product could receive was 2 and the highest was 12.

The GIM score attempts to rate compositions in a general, global manner (See Cooper & Odell, 1977 for a complete discussion). Covered by the GIM are any qualities that impress the rater including the quality of thought evidenced in the paper, the effectiveness of the overall organization, the effectiveness of the overall presentation, and the effectiveness of the style and word choice.

Analytic Scales

Like General Impression Marking, the Analytic Scales are quick, qualitative judgements about the effectiveness of a piece of writing. Analytic Scales, however, are measures of the effectiveness of particular qualities of a paper. For the provincial writing assessment, nine different qualities were

selected for testing, five rhetorical or discourse categories, four mechanics and convention categories. These scales were developed because they seemed to be the qualities most often evaluated in writing by teachers and because they were objectives most often identified as being taught by teachers.

Like the GIM, a paper was assigned a score between 1 and 6 (1 - low, 6 - high), but for each of the nine qualities.

Rhetorical Categories

1. Organization and Sequence
2. Wording and Style
3. Focus or Theme
4. Ideas
5. Choice of Details and Elaboration of Ideas

Mechanics and Conventions Categories

6. Sentence Structure
7. Usage
8. Punctuation and Capitalization
9. Spelling

A scorer, usually different from the GIM scorers, read each written product for the Analytic Scales. Specific descriptions of the qualities associated with high, middle, and low papers are given in each grade level Manual and Scoring Key, as well as examples of papers. (For a more complete discussion of analytic scales in general, refer to Diederich, 1974, and Cooper & Odell, 1977.)

For the provincial sample, every paper was rated on all nine analytic scales by at least one rater. A subset of papers at each grade level was marked analytically by a second rater in order to calculate the amount of agreement or reliability between independent raters.

These two scoring procedures (GIM and Analytic) yielded ten different scores for each written product. At the grade 3, 6, and 9 levels, students wrote two pieces of extended writing so that each student in those grades received 20 different scores (10 for each writing product). At the grade 12 level, the students were asked to write one piece of extended writing so that each student received 10 different scores. This is summarized in Table 2 (p. 136).

The GIM and Analytic score results for students in English 301 are based on scoring the entire 301 provincial sample of 311 test papers.

TABLE 2

GIM AND ANALYTIC SCORING OF WRITTEN PRODUCTS

Scales	Narrative Writing 3	6	9	12	Descriptive/Expository Writing 3	6	9	12
GIM (Scores range from 2 to 12)	x	x	x	-	x	x	x	x
Analytic (Scores range from 1 to 6)								
1. Organization and sequence	x	x	x	-	x	x	x	x
2. Wording and style (Word Choice)	x	x	x	-	x	x	x	x
3. Focus or theme	x	x	x	-	x	x	x	x
4. Ideas	x	x	x	-	x	x	x	x
5. Choice of Details & Elaboration of Ideas	x	x	x	-	x	x	x	x
6. Sentence Structure	x	x	x	-	x	x	x	x
7. Usage	x	x	x	-	x	x	x	x
8. Punctuation & Capitalization	x	x	x	-	x	x	x	x
9. Spelling	x	x	x	-	x	x	x	x

Essay Writing: GIM Results

Students at the grade 12 level were asked to write an essay explaining how something is done (Part IV of the test). Topics were suggested, such as: How to shop for a used car, How to train for your favourite sport, How to help a person in distress, etc. The criteria used in the Analytic scales were also given. Students were given approximately 60 minutes to write their stories.

The results of the GIM scoring for the essay are presented on page 137. Scores for the two scorers combined ranged from 2 to 12, with scores, 2, 3, and 4 described as "Low"; 5, 6, 7, 8, and 9 described as "Middle"; and 10, 11, and 12 described as "High".

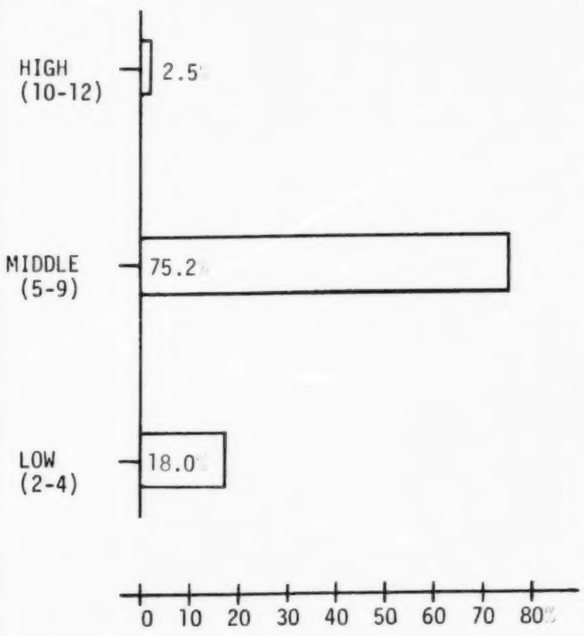
The results are presented in three ways:

1. The percentage of students obtaining each GIM score: e.g., 0.6% of English 301 students received a combined GIM score of 12 (both scorers awarded a 6 to these stories).
2. The percentage of students obtaining GIM scores categorized as High, Middle, and Low: e.g., 2.5% of 301 students received combined GIM scores of 10, 11, or 12 (High).
3. The average (mean) performance of students on the GIM scale (and the standard deviation). These are based on the scale of 2 to 12.

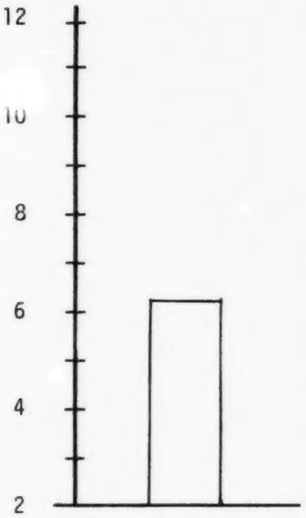
ESSAY WRITING: 1982 GIM RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
	12	0.6
HIGH	11	0.6
	10	1.3
	9	5.8
	8	12.2
MIDDLE	7	21.5
	6	17.7
	5	18.0
	4	9.3
LOW	3	6.4
	2	2.3
Insufficient Information		4.2

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 6.12

Standard Deviation = 1.86

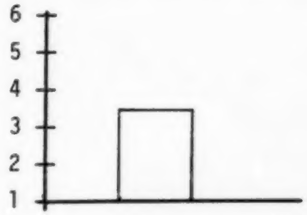
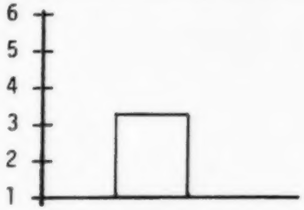
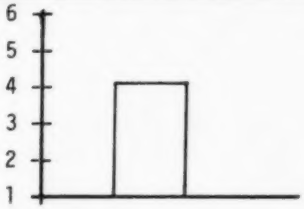
Essay Writing: Analytic Scale Results

The results of the analytic scoring for the grade 12 essay are presented on pages 139 to 141. All of the scores range on a scale from 1 to 6. Scores 1 and 2 are described as "Low", 3 and 4 as "Middle", and 5 and 6 as "High".

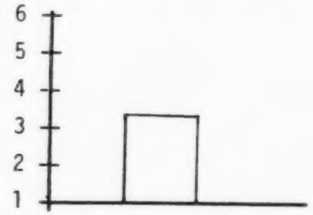
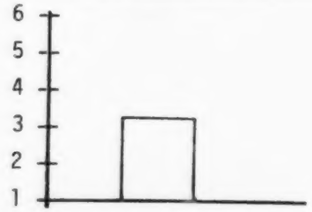
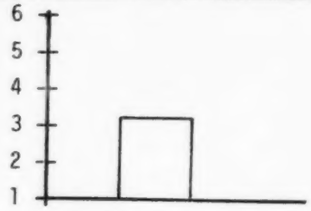
Results are presented in two ways:

1. Percentage of students obtaining each score: e.g., 1.9% of 301 students obtained a score of 6 on the analytic scale Organization and Sequence.
2. Average (mean) performance of students, and standard deviation, based on the scale of 1 to 6: e.g., the average of 301 students was 3.46 (standard deviation was 1.09) on Organization and Sequence.

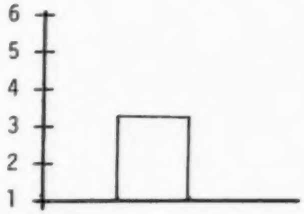
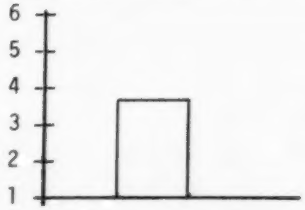
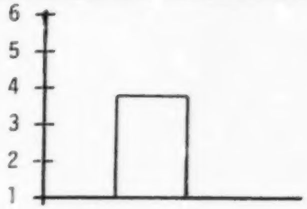
ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 301 STUDENTS

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
1. Organization and Sequence	HIGH	6	1.9%	
		5	14.1	
	MIDDLE	4	32.5	
		3	31.2	
	LOW	2	13.2	
		1	4.2	
	Insufficient Information		2.9	Average = 3.46 Standard Deviation = 1.09
2. Wording and Style	HIGH	6	0.3	
		5	7.7	
	MIDDLE	4	29.3	
		3	37.9	
	LOW	2	18.3	
		1	3.9	
	Insufficient Information		2.6	Average = 3.20 Standard Deviation = 0.98
3. Focus or Theme	HIGH	6	6.1	
		5	33.1	
	MIDDLE	4	29.9	
		3	20.6	
	LOW	2	7.1	
		1	0.6	
	Insufficient Information		2.6	Average = 4.09 Standard Deviation = 1.07

ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 301 STUDENTS

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
4. Ideas	HIGH	6	1.6	
		5	9.3	
	MIDDLE	4	33.8	
		3	36.0	
	LOW	2	14.1	
		1	2.6	
	Insufficient Information		2.6	Average = 3.39 Standard Deviation = 0.99
5. Choice of Details and Elaboration of Ideas	HIGH	6	0.3	
		5	10.6	
	MIDDLE	4	29.3	
		3	36.0	
	LOW	2	15.1	
		1	5.8	
	Insufficient Information		2.9	Average = 3.26 Standard Deviation = 1.05
6. Sentence Structure	HIGH	6	1.0	
		5	8.4	
	MIDDLE	4	26.0	
		3	37.0	
	LOW	2	20.6	
		1	4.5	
	Insufficient Information		2.6	Average = 3.17 Standard Deviation = 1.04

ESSAY WRITING: 1982 ANALYTIC SCALE RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Analytic Marking Scale			Percentage of Students Obtaining Each Score	Average Performance
Objective	Scale			
7. Usage	HIGH	6	1.0	 Average = 3.23 Standard Deviation = 1.14
		5	14.5	
	MIDDLE	4	21.5	
		3	35.4	
	LOW	2	19.0	
		1	6.1	
	Insufficient Information		2.6	
8. Punctuation and Capitalization	HIGH	6	4.5	 Average = 3.69 Standard Deviation = 1.20
		5	21.5	
	MIDDLE	4	32.2	
		3	21.2	
	LOW	2	14.8	
		1	3.2	
	Insufficient Information		2.6	
9. Spelling	HIGH	6	7.7	 Average = 3.77 Standard Deviation = 1.34
		5	24.1	
	MIDDLE	4	24.8	
		3	21.5	
	LOW	2	10.9	
		1	6.1	
	Insufficient Information		4.8	

WRITTEN PRODUCT:

ATTRIBUTE AND DESCRIPTIVE SCALES AND ERROR COUNTS

In addition to the qualitative judgements made on each paper using the GIM and Analytic Scales, a randomly selected sub-sample of papers was examined for more intensive structural, linguistic, attribute, and error descriptions. These in-depth analyses were carried out by the same group of raters employed for the other marking tasks. A total of 55 student papers were selected at the grade 12 level for this intensive scoring.

The rationale for these types of analyses was to make assessments regarding the kinds of structures students used in their actual writing and the kinds of errors most frequently made by students in actual writing situations. Also, they were conducted in order to draw some conclusions about the number of students demonstrating serious problems in the structural aspects of their writing. The choice of structures, errors, and attributes was based on the survey of writing done by Hunt (See Cooper & Odell, 1977), on the work of Evans (1979) in the Ontario assessment, and on the objectives identified by practicing teachers as the ones of greatest interest and use to them in their classrooms.

Attribute Marking Scales

Seven different attribute marking scales were identified as being of interest to teachers. The scales themselves were evaluations about whether or not a student was experiencing serious (in need of remedial help) difficulty in particular areas. For each paper evaluated, the rater was asked to rate the composition on the seven attributes, each on a three-point scale:

3 - No errors;

2 - An occasional error, but not rated as a serious problem

1 - Serious problem (Student is in need of special remedial help);

X - Insufficient information (Structures not used or insufficient writing for an evaluation to be made).

Five of the seven areas were evaluated at all grade levels. The last two (use of contractions and use of quotation marks in direct discourse) were evaluated only at grade levels 3 and 6. The seven attributes were:

1. Run-on sentences,
2. Use of complete sentences,
3. Paragraphing,
4. Internal punctuation,
5. Capitalization,
6. Use of contractions, and
7. Use of quotation marks in direct discourse.

Descriptive Scales

The descriptive scales were counts of particular structures used by students. The following counts were made for each student composition in grades 3, 6, and 9. Grade 12 compositions were tallied on all but the last one

1. Total number of words written
2. Total number of T-units (independent clauses) written
3. Total number of dependent clauses written
4. Total number of adverbial clauses written

5. Total number of relative clauses written
6. Total number of noun clauses written
7. Total number of modifiers used (not counted)
8. Total number of instances of direct discourse (not counted)

The particular counts were chosen because they seem closely related both to composition quality and to syntactic fluency (See Hunt in Cooper & Odell, 1977). The purpose was to judge whether or not students easily and fluently used these structures in their actual writing.

From these counts, a number of transformed scores were calculated to produce twelve indices of writing maturity in students. In general, the more fluent, mature writers will demonstrate higher scores on each of these indices, whereas weaker, less mature writers will score lower. These scores should be viewed, however, only as indicators of how students handle the syntax and grammar of the language in free writing situations. The question answered by these data is: Do students use these particular structures in their own writing? The question of whether or not students can use these structures is left to be answered by the sentence writing tasks (Part III).

The scores calculated included:

Words per T-unit
Words per clause
Clauses per T-unit
Dependent clauses per T-unit
Adverbial clauses per T-unit
Adverbial clauses per 100 words
Relative clauses per T-unit
Relative clauses per 100 words
Noun clauses per T-unit
Noun clauses per 100 words
Direct discourse per T-unit
Direct discourse per 100 words

Error Counts

Error counts, though tedious to do, can give information about how well students avoid error in actual writing situations. These counts are realistic alternatives to the use of objective (particularly multiple-choice type) items in evaluating writing. Error counts can give greater information about the choices that students are making in actual writing tasks.

Each of the error categories are calculated in these ways:

average number of errors per composition
average number of errors per T-unit (i.e. per independent clause), and
average number of errors per 100 running words

The error counts obtained were:

1. Number of sentence fragments
2. Number of incorrectly placed modifiers
3. Number of sentences not beginning with a capital letter
4. Number of proper nouns not capitalized
5. Number of sentences not ending with appropriate punctuation
6. Number of internal comma errors
7. Number of faulty subject-verb agreement
8. Number of verb tense errors
9. Number of word order errors

Similarly to the descriptive scales a number of scores were calculated from the error counts. But, opposite to the descriptive scales, higher scores on the scale reflect greater error rates.

The calculated scores were:

- Sentence fragments per T-unit
- Sentence fragments per 100 words
- Incorrectly placed modifiers per T-unit
- Incorrectly placed modifiers per 100 words
- Beginning capitalization errors per T-unit
- Beginning capitalization errors per 100 words
- Proper nouns not capitalized per T-unit
- Proper nouns not capitalized per 100 words
- End punctuation errors per T-unit
- End punctuation errors per 100 words
- Comma errors per T-unit
- Comma errors per 100 words
- Faulty agreement per T-unit
- Faulty agreement per 100 words
- Verb tense errors per T-unit
- Verb tense errors per 100 words
- Word order errors per T-unit
- Word order errors per 100 words

Story Writing: Attribute and Descriptive Scales/Error Counts

The results for the essays of the English 301 students (Part IV) are presented for:

- The 7 attribute scales (p. 145),
- The 8 descriptive scales (p. 146).
- The 12 indices calculated from the descriptive scales (p. 146),
- The 9 error counts (p. 147),
- The 18 indices calculated from the error counts (p. 147).

Since there are a large number of scales, the results have been tabulated in a compacted form in the following tables.

The average (mean) and standard deviation are reported for each scale and index. For the attribute scales, the percentage of students obtaining each score is reported: e.g., 17.4% of students obtained a score of 3 (no errors) on the attribute scale "Run-on sentences", and the average is 1.96 on the scale of 1 to 3 (standard deviation is 0.64).

Of the 55 essays intensively scored at the grade 12 level, 23 are from English 301 students. The results are based on these 23 essays.

ESSAY WRITING: 1982 ATTRIBUTE SCALES RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
1. Run-on sentences				
No errors	3	17.4		
An occasional error	2	60.9	1.96	0.64
Serious problem	1	21.7		
Insufficient information		0.0		
2. Use of complete sentences				
No errors	3	52.2		
An occasional error	2	26.1	2.30	0.82
Serious problem	1	21.7		
Insufficient information		0.0		
3. Paragraphing				
No errors	3	30.4		
An occasional error	2	52.2	2.18	0.66
Serious problem	1	13.0		
Insufficient information		4.3		
4. Internal punctuation				
No errors	3	0.0		
An occasional error	2	52.2	1.52	0.51
Serious problem	1	47.8		
Insufficient information		0.0		

Attribute Marking Scale		Percentage of Students Obtaining Each Score	Average Performance Mean	Average Performance Standard Deviation
5. Capitalization				
No errors	3	43.5		
An occasional error	2	47.8	2.41	0.59
Serious problem	1	4.3		
Insufficient information		4.3		
6. Use of contractions				
No errors	3			
An occasional error	2	not counted		
Serious problem	1			
Insufficient information				
7. Use of quotation marks in direct discourse				
No errors	3			
An occasional error	2	not counted		
Serious problem	1			
Insufficient information				

ESSAY WRITING: DESCRIPTIVE SCALE RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Descriptive Scale (Actual counts)	Average Number	
	Mean	Standard Deviation
1. Words	483.39	170.18
2. T-units	36.22	15.11
3. Dependent Clauses	17.26	8.20
4. Adverbial clauses	10.04	4.76
5. Relative clauses	3.00	2.52
6. Noun clauses	4.26	4.04
7. Modifiers	(not counted)	
8. Instances of direct discourse	(not counted)	

Indices calculated from Descriptive Scales	Mean	Standard Deviation
Words per T-unit	13.74	2.98
Words per clause	9.10	1.25
Clauses per T-unit	1.51	0.27
Dependent clauses per T-unit	0.51	0.28
Adverbial clauses per T-unit	0.30	0.16
Adverbial clauses per 100 words	2.16	0.93
Relative clauses per T-unit	0.094	0.098
Relative clauses per 100 words	0.65	0.54
Noun clauses per T-unit	0.12	0.11
Noun clauses per 100 words	0.83	0.69
Direct discourse per T-unit	-	-
Direct discourse per 100 words	-	-

ESSAY WRITING: ERROR COUNTS FOR ENGLISH 301 STUDENTS (Grade 12)

Indices Calculated From Error Counts	Mean	Standard Deviation
Sentence fragments per T-unit	0.031	0.057
Sentence fragments per 100 words	0.21	0.34
Incorrectly placed modifiers per T-unit	0.032	0.047
Incorrectly placed modifiers per 100 words	0.23	0.35
Beginning capitalization errors per T-unit	0.076	0.091
Beginning capitalization errors per 100 words	0.59	0.72
Proper nouns not capitalized per T-unit	0.0	-
Proper nouns not capitalized per 100 words	0.0	-
End punctuation errors per T-unit	0.74	0.076
End punctuation errors per 100 words	0.56	0.58
Comma errors per T-unit	0.24	0.15
Comma errors per 100 words	1.81	1.22
Faulty agreement per T-unit	0.017	0.035
Faulty agreement per 100 words	0.13	0.25
Verb tense errors per T-unit	0.040	0.067
Verb tense errors per 100 words	0.28	0.45
Word order errors per T-unit	0.044	0.065
Word order errors per 100 words	0.32	0.48

Error Counts		Percentage of Students Obtaining each Count	Average Error Count Mean	Average Error Count Standard Deviation
1. Sentence fragments	0	60.9%	0.87	1.39
	1	17.4		
	2 or more	21.7		
2. Incorrectly placed modifiers	0	56.5	0.96	1.33
	1	13.0		
	2 or more	30.5		
3. Sentences not beginning with Capital letters	0	30.4	2.48	2.92
	1	21.7		
	2 or more	47.9		
4. Proper nouns not capitalized	0	95.7	-	-
	1	4.3		
	2 or more	0.0		
5. Sentences not ending with appropriate punctuation	0	26.1	2.39	2.69
	1	17.4		
	2 or more	56.5		
6. Internal comma errors	0	4.3	8.65	6.32
	1	4.3		
	2 or more	91.4		
7. Instances of faulty subject/verb agreement	0	69.6	0.52	0.90
	1	13.0		
	2 or more	17.3		
8. Verb tense errors	0	56.5	1.48	2.33
	1	8.7		
	2 or more	34.8		
9. Word order errors	0	56.5	1.22	1.65
	1	4.3		
	2 or more	39.2		

ESSAY PLANNING

The purpose of the discourse planning task at the grade 12 level was to assess students' ability in pre-planning an essay. Since pre-writing and planning is an important aspect of the act of composing, it was decided that students at grade 12 could be asked to plan a persuasive essay. The topic was whether students in high school should hold part-time jobs.

In the task itself, students were given a list of ideas for the essay. Both important and unimportant ideas on both sides of the issue were interspersed. Students were then asked to plan an essay by identifying a title, a major theme, and a number of major ideas supporting that theme, and then to group important details under each of the major ideas. They were also asked to write the concluding paragraph of the essay.

The essay plan was rated using three analytic marking scales; these were indicated to students:

A. Organization and sequence,
B. Focus or theme, and
C. Choice of details and elaboration.

Raters read the essay plans, then reacted to them holistically by assigning a score of 1 to 6 (1 - low, 6 - high) for each of the three qualities. Particular attributes of high, middle, and low plans are described in the grade 12 Manual and Scoring Key.

All essay plans were rated by two different scorers in much the same manner as was done for the general impression marking of the essay (see p. 134). A third scorer was assigned to arbitrate the essay plan if the first two scorers differed by two or more points on two or three of the scales.

The scores of the first two scorers were combined to form a scale that ranged from 2 (1 + 1) to 12 (6 + 6). In the arbitrated cases the score was calculated by:

$$\frac{\text{FIRST SCORE + SECOND SCORE}}{2} + \text{ARBITRATOR SCORE} + \text{COMBINED SCORE}$$

Essay Planning: Results

Teachers were instructed to read the directions aloud to students for the essay planning task (Part II). Students were given approximately 45 minutes to complete their plan. These plans were rated on the three scales indicated above.

The results of the 311 English 301 papers are presented on pages 149 to 151 in these three ways:

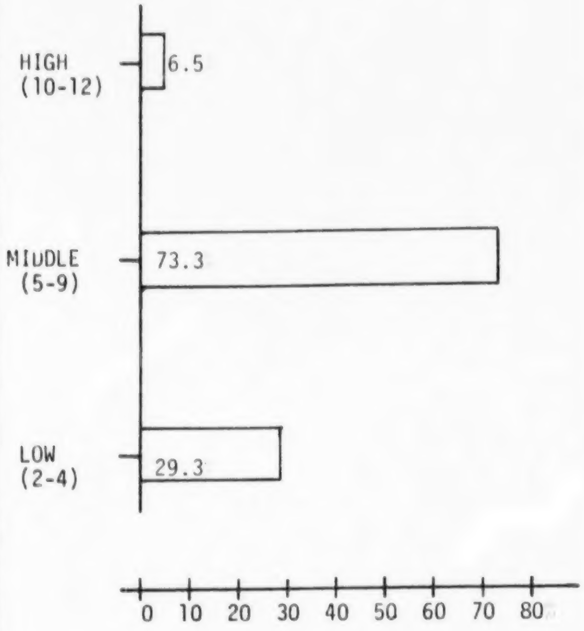
1. The percentage of students obtaining each combined score: e.g., on the scale "Organization and Sequence", 0.0 of English 301 students received a combined score of 12 (both scorers rated it a 6).
2. The percentage of students obtaining combined scores categorized as High (10-12), Middle (5-9), and Low (2-4): on "Organization and Sequence", 6.5 of English 301 students were in "High" category.
3. The average (mean) performance of students on each scale (and standard deviation). These are based on the scale of 2 to 12.

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

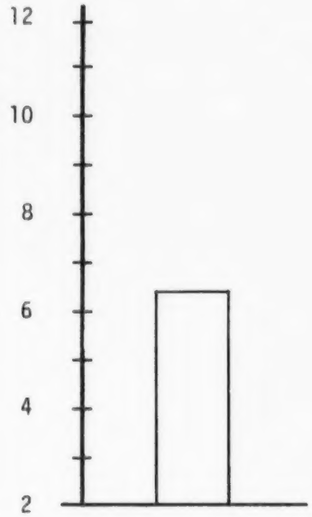
A. ORGANIZATION AND SEQUENCE

GIM Scale (two scorers (combined)		Percentage of 1982 Students Obtaining Each Score
HIGH	12	0.0
	11	2.3
	10	4.2
MIDDLE	9	10.0
	8	9.6
	7	19.6
	6	16.7
	5	17.4
LOW	4	12.2
	3	5.5
	2	1.6
Insufficient Information		0.9

Percentage of 1982 Students Obtaining Each Score Category



Average Performance

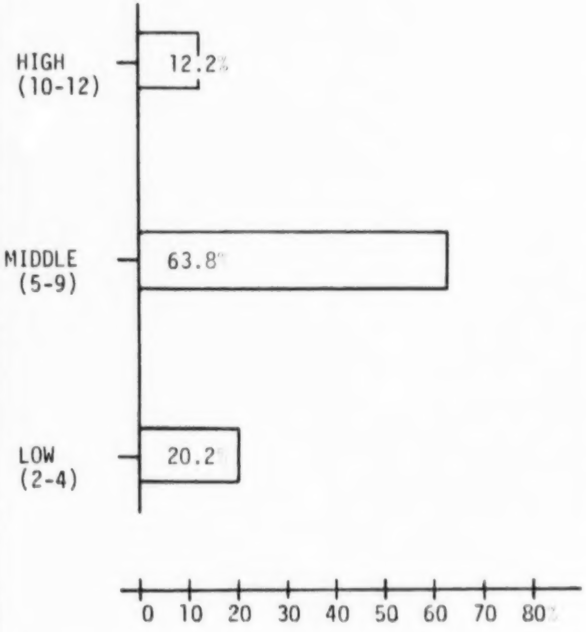
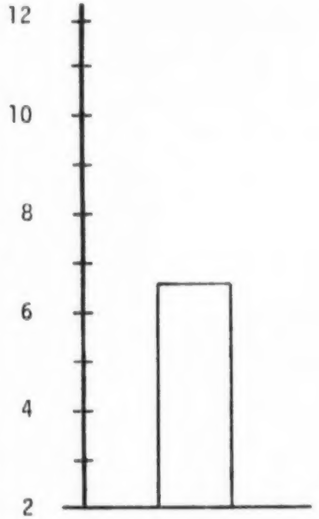


Average = 6.35

Standard Deviation = 2.02

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

B. FOCUS OR THEME

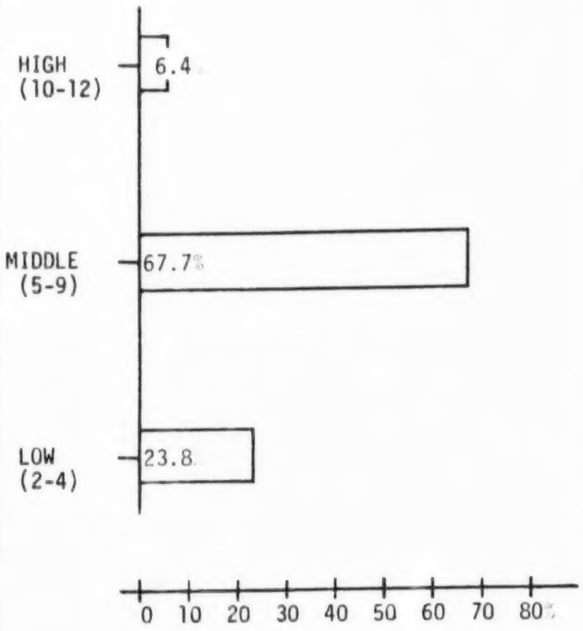
GIM Scale (two scorers (combined)		Percentage of 1982 Students		Average Performance
		Obtaining Each Score	Obtaining Each Score Category	
	12	0.3%		
HIGH	11	3.5		
	10	8.4		
	9	9.3		
	8	9.9		
MIDDLE	7	14.8		
	6	14.1		
	5	15.7		
	4	12.2		
LOW	3	6.1		
	2	1.9		
Insufficient Information		3.5		Average = 6.54 Standard Deviation = 2.27

ESSAY PLANNING: 1982 RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

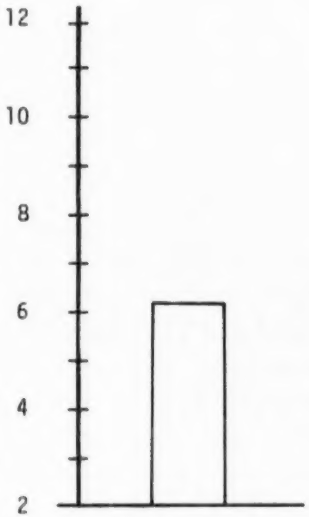
C. CHOICE OF DETAIL AND ELABORATION

GIM Scale (two scorers (combined)		Percentage of 1982 Students: Obtaining Each Score
HIGH	12	0.3%
	11	1.6
	10	4.5
MIDDLE	9	6.4
	8	12.2
	7	19.2
	6	13.5
	5	16.4
LOW	4	11.9
	3	9.3
	2	2.6
Insufficient Information		1.9

Percentage of 1982 Students Obtaining Each Score Category



Average Performance



Average = 6.16

Standard Deviation = 2.10

SPELLING

The purpose of the spelling assessment on the 1982 Writing Assessment Program was to determine if students could spell correctly the words they actually used in their writing as well as the words they were likely to need in their writing. Spelling in the actual written product at the Grade 12 level was assessed by one of the nine analytic scales (See Analytic Scales). In order to assess the performance of students on the spelling of words they were likely to need in their general writing (but may not have used in the writing tasks provided on the test), words were randomly selected from the vocabulary lists by Dale and O'Rourke (1976). This list of words assigns words to different grade levels based on vocabulary recognition tests generally given in North America. It was chosen because no recent survey of words that Grade 12 students use in general writing was available, either for Canada or the United States.

Unlike the 1979 assessment, the 1982 Manitoba Writing Assessment Program tests spelling by dictation (rather than by error recognition and correction in a multiple-choice format). This decision was based on the conclusions regarding the most valid form of testing spelling drawn by Thomas (1979). Unfortunately, because of the difference in test format, realistic comparisons between 1979 spelling performance and 1982 spelling performance at the Grade 12 level cannot be made. The words chosen for the 1982 assessment, however, seem more reflective of the words students need in general writing than those selected for the 1979 assessment.

There is no breakdown of the spelling words into categories. Thus, the results are based on the number of correctly spelled words out of the total of 20.

Reporting of the 1982 Results

The average (mean) score is reported for the scores based on the 20 words. A frequency distribution indicating the percentage of students obtaining each score is presented (e.g., 3.9% of English 300 students spelled 19 or 20 of the words correctly). Finally, the percentage of students spelling each word correctly is indicated in a separate table.

The results for the 311 English 301 students are given on page 153.

SUMMARY OF THE 1982 SPELLING RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Score Correct out of Total Number of Words (25)	Percentage of Students Obtaining Each Score
19-20	3.9%
17-18	13.2
15-16	19.0
13-14	21.2
11-12	20.2
9-10	10.0
7-8	6.7
5-6	4.2
3-4	1.3
1-2	0.0
0	0.3

Words Selected For the 1982 Assessment	
Items	Percentage of 1982 students obtaining item correct
1. basically	83.3%
2. possessed	52.1
3. prescriptive	84.9
4. wheelbase	96.8
5. urchin	41.2
6. bolster	73.6
7. winsome	45.3
8. nimble	84.6
9. deluge	64.6
10. nectarine	51.4
11. incapacitate	49.2
12. ungrudging	73.3
13. decrepit	33.4
14. croon	42.8
15. collective	96.1
16. slipshod	52.1
17. sedimentary	68.2
18. boycott	64.0
19. entreat	83.3
20. itinerary	50.5

Average Score Correct	12.9
Standard Deviation	3.61
Average Percentage Correct	64.5%

SENTENCE WRITING

A number of sentence writing tasks were designed to assess performance on selected structural, syntactic, and usage objectives. The particular format of the sentence writing tasks was chosen for five reasons:

> It closely resembled the kind of thinking and shaping a writer probably goes through when composing.
>
> It allowed for a variety of answers which could be labeled "correct" so that a student was not limited to a single correct answer and could create the most effective sentence(s) for the situation and information provided.
>
> The structurally correct answers in most cases could be generally categorized from least mature (or fluent) to most mature (or fluent).
>
> The open format centred on what students could do in particular circumstances rather than on what they could not do.
>
> The format could allow for the assessment of whether students could use particular structures.

Performance on the sentence writing tasks could be compared with the descriptive scales in the actual writing tasks in order to determine if students in fact did use the targeted structures in free writing.

Although a number of objectives (or targeted structures) were common across all four grade levels, the particular structures or conventions assessed at each grade level were based on objectives that practicing teachers felt were applicable at that level.

Objectives at the Grade 12 Level

Objectives at the grade 12 level are listed below; the sections comprised Part III of the student test.

Section A - Objective: Use of participles
Section B - Objective: Use of parallel structure
Section C - Objective: Use of relativization
Section D - Objective: Use of adverbial subordination

Several sentences were required for each section to provide a more accurate indicator of student ability.

The objectives fall into one major area: syntactic maturity. All students should have some understanding of, as well as considerable skill in, this area.

Students' sentences were scored according to several categories; these were denoted a, b, c, etc. It is likely that the most careful, mature writers would respond consistently with answers that would fall into category "a" of each objective. Less mature writers might have a profile of mixed performance (e.g., use participles correctly and form sentences with adverbial subordination but not consistently use parallel structure or relativization). The fact that a student did not necessarily mean that she/he does not know how to employ the structure or convention, rather that the structure or convention is not readily accessible in the written mode. The very weakest students might demonstrate little fluency in the structures they write or choose. The tasks presented in the assessment allow for divergent answers so that the teacher can assess the level of maturity or fluency of students.

An Example

For Item 7 of Section D of the grade 12 test the student was asked to combine the following two sentences using one of the words suggested.

7. There was an excited and cheerful atmosphere in the gym. The home team had won the basketball game.

The most fluent writer would probably have responded with an answer such as:

> There was an excited and cheerful atmosphere in the gym because the home team had won the basketball game.

This answer would be categorized a: Combined using adverbial subordination.

A somewhat less fluent writer, however might have responded with an answer such as:

> There was an excited and cheerful atmosphere in the gym and the home team had won the basketball game.

This answer would be categorized b: Correctly combined using only coordination. This answer is structurally correct and grammatical; it is simply not as efficient or fluent as the first answer. The categorization scheme for this objective, then, allows the teacher to judge a structure as "correct", but "less mature".

Another writer might have responded with:

> The home team having won the basketball game, there was an excited and cheerful atmosphere in the gym.

By no means is this a less mature answer than the one categorized as "a"; however, it is different enough to warrent being categorized separately. Since it was impossible to anticipate all "correct" answers, a category, c: Correctly combined in some other fashion, was devised to cover the answers not covered by "a" or "b". Here the student has written a structurally correct, grammatical, and meaningful sentence and should be given credit for it, even though it did not reflect the objective that we as test writers had in mind. The maturity or fluency level is not apparent from the category description; the above answer might be categorized with a somewhat less mature answer. It is included in order to give the writer credit for meaningful writing.

The final category, d: Incorrectly combined, was used only if the student wrote an answer that did not make sense or was not grammatical.

The categories for scoring varied according to the particular objective, but in all cases the scheme was similar to the example presented. The complete categories and examples of answers are given in the grade 12 Manual and Scoring Key.

Results for the Sentence Writing

At the grade 12 level students were asked to do two or three writing tasks (items) for each objective. The results are presented for each sentence and for each objective. Since the scoring categories for each objective were not based on a numerical scale averages cannot be used. Instead reported are the percentages of students whose sentences were categorized in each of the scoring categories (the first scoring category is always the targeted structure of the objective). Also reported is the average of the percentages for a particular category for the two items pertaining to an objective. For example, in Section A, 86.5% of English 301 student sentences were categorized as a: Used participles for Item 1 and 85.5% for Item 2. On average, 86.0% of student sentences were categorized a: Used participles for the two sentences combined, and thus for the objective: Use of participles.

The results are based on the total English 301 sample of 311 test papers.

SENTENCE WRITING: 1982 RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items
Section A-Objectives: Use of participles			
	1. The quarterback leaned back and threw the pass. The quarterback was standing in the clear. The pass was the winning one.	2. The hunters pushed the canoe out into the river and began to paddle toward the opposite shore. The hunters were keeping an eye out for deer. The river was rushing.	
a - Participles used	86.5	85.5	86.0
b - Correctly combined in some other fashion	4.8	5.1	5.0
c - Incorrectly combined	8.0	8.0	8.0
Omitted	0.6	1.2	0.9
Section B-Objective: Use of parallel structure			
	3. Harold couldn't remember where he sat in class. Harold couldn't remember whom he sat beside in class. Harold couldn't even remember the classroom number.	4. Centuries ago people thought the world was flat. They also thought that the oceans were filled with huge monsters. People thought if they sailed too far away from land, they would fall off the world.	
a - Correctly combined with structure parallel	58.2	46.9	52.6
b - Correctly combined in some fashion not requiring parallelism	10.0	13.8	11.9
c - Structures not parallel	24.4	24.1	24.3
d - Incorrectly combined	6.4	14.1	10.3
Omitted	1.0	1.0	1.0

SENTENCE WRITING: 1982 RESULTS FOR ENGLISH 301 STUDENTS (Grade 12)

Scale Categories	Percentage of students obtaining each category for the Items		Percentage of students obtaining each category: Average percentage for the items	
Section A-Objective: Use of participles				
	1. The quarterback leaned back and threw the pass. The quarterback was standing in the clear. The pass was the winning one.	2. The hunters pushed the canoe out into the river and began to paddle toward the opposite shore. The hunters were keeping an eye out for deer. The river was rushing.		
a - Participles used	81.7%	84.2%	83.0%	
b - Correctly combined in some other fashion	1.0	1.0	1.0	
c - Incorrectly combined	14.5	11.3	12.9	
Omitted	2.9	3.6	3.3	
Section B-Objective: Use of parallel structure				
	3. Harold couldn't remember where he sat in class. Harold couldn't remember whom he sat beside in class. Harold couldn't even remember the classroom number.	4. Centuries ago people thought the world was flat. They also thought that the oceans were filled with huge monsters. People thought if they sailed too far away from land, they would fall off the world.		
a - Correctly combined with structure parallel	72.3%	70.1%	74.6%	72.3%
b - Correctly combined in some fashion not requiring parallelism	9.6	8.0	5.5	7.7
c - Structures not parallel	10.6	8.4	8.0	9.0
d - Incorrectly combined	5.5	11.6	9.3	8.8
Omitted	1.9	1.9	2.6	2.1